Project no. 727040

GIFT
Meaningful Personalization of Hybrid Virtual Museum Experiences Through Gifting and Appropriation

Horizon 2020

SC6-CULT-COOP-2016-2017

CULT-COOP-08-2016
Virtual museums and social platform on European digital heritage, memory, identity and cultural interaction.

Start date: 1 January 2017. Duration: 36 months

D3.4
Report on Sensitive Pictures

Due date: 31 December 2019

Actual submission date: 20 January 2020

Number of pages: 50

Lead beneficiary: IT University Copenhagen

Author(s): Anders Sundnes Løvlie, Edgar Bodiaj, Karin Ryding, Steve Benford, Harriet Cameron, Dimitrios Darzentas, Christian Hviid Mortensen, Paulina Rajkowska, Carolina Fuentes, Velvet Spors, Bogdan Spanjevic

Project Consortium

Beneficiary no.	Beneficiary name	Short name
1 (Coordinator)	**IT University of Copenhagen**	**ITU**
2	Blast Theory	Blast Theory
3	Next Game	NextGame
4	University of Nottingham	UoN
5	Uppsala University	UU
6	Europeana Foundation	EF
7	Culture24	C24

Dissemination Level

PU	Public	X
CO	Confidential, only for members of the consortium (including the Commission Services)	
EU-RES	Classified Information: RESTREINT UE (Commission Decision 2005/444/EC)	
EU-CON	Classified Information: CONFIDENTIEL UE (Commission Decision 2005/444/EC)	
EU-SEC	Classified Information: SECRET UE (Commission Decision 2005/444/EC)	

Type

R	Document, report	X
DEM	Demonstrator, pilot, prototype	
DEC	Websites, patent filling, videos, etc.	
O	Other	
ETHICS	Ethics requirement	

Table of contents

1 Introduction **4**

2 Background/related work **5**

2.1 Emotions in computing and design 6

2.2 Affective experiences in museums 7

3 Approach **8**

4 Sensitive Pictures **10**

4.1 Experience design 10

4.2 Technical implementation 16

5 Test results - August 2019 **19**

5.1 The user experience 19

Overall reception 19

The six themes represent a large variety of personal experiences 20

Describing emotions are challenging 21

People try their best to make sense of emotion data 22

Skepticism and interest to the idea of computers trying to read emotions 22

5.2 Quantitative analysis of interview data 23

Emotions 26

5.3 Personal data and trust 28

5.4 Analysis of emotion detection data 34

5.5 Visualisations from Emotion Mapper 37

6 Sensitive Pictures with EEG headsets **45**

6.1 EEG prototype 45

6.2 Data visualisations from the EEG prototype 45

7 Concluding remarks **49**

References **50**

1 Introduction

The topic of this report is a study of "Sensitive Pictures", the museum experience presented in deliverable D3.3. The purpose of this work is twofold: On the one hand, to explore the design of an emotionally engaging museum visitor experience; and on the other, to explore ways to facilitate emotional responses and a dialogue with visitors about emotions.

The study was fittingly situated in the museum of the Norwegian artist Edvard Munch (1863-1944), famous for his dynamic exploration of emotions. Munch's best known work, The Scream, has become one of the most iconic images of world art. It has been widely interpreted as representing the universal anxiety of modern man (Eggum 1984). As can be seen in The Scream, Munch actively worked with translating emotional states into highly striking works of art.

A hybrid museum experience was designed in collaboration between representatives of the museum, and teams from the GIFT partners NextGame, the University of Nottingham (UoN) and the IT University of Copenhagen (ITU). The design allowed us to explore several different ways to provoke emotional responses in museum visitors and to collect information about these responses from them. We enhanced the viewing of paintings in the museum with audio commentary that was intended to provoke emotional responses. We then employed a combination of techniques for capturing information relating to visitors' emotional responses: (i) self-reporting of emotions using the Self-Assessment Manikin (SAM) model; (ii) self-reporting thoughts and feelings via free text answers to questions; and (iii) using video analysis of facial movements and electroencephalogram (EEG) sensors to capture physiological signals pertaining to emotional responses. We then utilised the Emotion Mapper tool from WP6 to create a series of visualisations that revealed and cross-related this information across combinations of paintings and participants as a way of stimulating reflection on emotional responses to museum artefacts.

By combining several different methods for capturing emotional states during an emotionally charged experience at the museum, as well as conducting semi-structured interviews with participants, we were seeking to answer the following overarching research question:

RQ: How can we explore emotional engagement in encounters with a museum exhibition, and facilitate a dialogue with visitors about emotions?

As a part of this, we explored the following sub-questions:

- How can visitors' emotions be captured as part of a hybrid museum experience? How can this be supported by different techniques, including self-reporting, emotion detection with computer vision, and EEG sensing?

- How do visitors experience having their emotions gauged using these techniques? How does this affect aspects of the experience such as enjoyment, reflection and trust?
- How can data collected with the above mentioned techniques be presented back to visitors, in order to engage them in dialogue about their emotional responses?

2 Background/related work

The question of 'what are emotions?' is both deeply fascinating and also profoundly challenging, and has engaged thinkers from across many different disciplines from neuroscience and psychology to the humanities and arts. These perspectives span different epistemologies, from those grounded in science that try to reduce emotion to complex interactions among the brain's various systems (neuroscience) or taxonomies of more or less fundamental emotions (psychology), to those that see emotions as something to be critically or artistically interpreted, reflected on or even provoked (arts and humanities). These complex relations have long been the subject of academic debate (e.g. Damásio 1995) but have been complicated further by the rise of computational and artificial intelligence-based techniques that aim to measure emotions through physiological signals (e.g. McStay 2018). The complex relationships between physiological signals and emotions and reflections and feelings is something that we wanted to explore in the following work.

Our interest in this project draws on and indeed deliberately juxtaposes these different traditions of thought about emotion. On the one hand we are fundamentally interested in working in and with museums, and in our specific case a museum of overtly emotionally-oriented fine art, that drives an interest in how visitors might interpret emotions in new ways, both in artworks and in themselves. On the other hand, we are considering how to employ digital technologies to achieve this. These technologies are often grounded in more scientific or reductionist models of emotion, because at the end of the day they need to implement algorithmic approaches to processing emotional data. Our interest here is to explore productive relationships between these two perspectives. How can new technologies that are concerned with 'measuring' emotion be employed in a setting that demands a thoughtful interpretation of emotion? In undertaking such an exploration, we are primarily concerned with creating experiences that provoke reflection among visitors, in part about the emotions expressed in the exhibited artworks, but also about the nature of the visitors' own emotions, and perhaps even on the role of technologies in detecting emotions within museum visits. We hope that as as a secondary contribution, our explorations may shed a little light onto attempts to measure and classify emotions as part of interactive and affective computing experiences.

2.1 Emotions in computing and design

Before 1990 emotions had a low-status as a topic of research and the focus was mainly on how emotions got in the way of rational thinking (Höök, 2012). This changed in the 90ies when a new wave of research on emotions and affect emerged with in the fields of psychology, neurology, medicine, and sociology (e.g. Damasio, 1995; Katz, 1999; Ledoux, 1996). This novel focus on emotions had a big impact on research and innovation of new technology. Within the fields of Human-Computer Interaction (HCI) and Interaction Design research on emotion and design went into three different directions with three very different theoretical perspectives (Höök, 2012).

The first direction is a cognitivistically inspired design approach named "Affective Computing" after a groundbreaking book by Rosalind Picard which came out in 1997. In her book, Picard implements a biologistic perspective on emotion processes. To her, Affective Computing describes computing that relates to, arises from or influences emotions (Picard, 1997, p. 1). The machine should interpret the emotional state of humans and adapt its behaviour to them, giving an appropriate response to those emotions. Methods used in this approach are quantitative and includes the use of sensors which capture data about the user's physical state or behaviour, for example the recording of facial expressions, body posture, and gestures. Other sensors detect emotional cues by directly measuring physiological data. An example of this is galvanic skin response (GSR), which refers to changes in sweat gland activity that are reflective of the intensity of our emotional state, otherwise known as emotional arousal (Sharma, Kacker, & Sharma, 2016).

The second direction in the research of design and emotions is a design approach called "Affective Interaction" and it can be seen as a counter-reaction to Affective Computing. It takes its inspiration from sociology and starts from a constructive, culturally-determined perspective on emotion. Affective Interaction draws upon phenomenology and sees emotion as constructed in interaction – between people and between people and machines (Boehner, DePaula, Dourish, & Sengers, 2005, 2007; Gaver, 2009; Höök, 2006; Höök, Ståhl, Sundström, & Laaksolahti, 2008; Sundström, Ståhl, & Höök, 2007). This approach forwards qualitative methods such as interviews, self-reporting of emotional states (Isbister, Höök, Sharp, & Laaksolahti, 2006), or cued-recall, a form of situated recall, as a method to elicit information about user affect during the use of a system (Bentley, Johnston, & von Baggo, 2005). As Boehner et al. points out,

> Sometimes emotions cannot be articulated by users in straightforward ways, yet informational approaches can unintentionally attempt to force users into a straightjacket of formalized expression. The interactional approach does not require emotion to be formalized by the system; instead, all the emotional meaning in the system can be supplied by the users. (2005, p. 66)

Finally, the third direction, which is less covered in this rapport, can be named "Technology as Experience", a design approach that refuses to singling out emotion from the overall interaction.

Instead, emotion can be seen as part of a larger whole of experiences we may design for (Gaver, 2009; McCarthy & Wright, 2007; Norman, 2004).

2.2 Affective experiences in museums

Museums have a unique ability to give visitors emotional, restorative and transformative experiences. This is made possible because museums are liminal spaces, which means that they are set apart from ordinary life (Bell, 2002). From visitor research it is known that a specific group of visitors, which are sometimes called 'Rechargers' (Falk and Dierking, 2012), primarily seek out these types of contemplative, spiritual and/or restorative experiences.

Moreover, research shows that emotions play an essential part in people's experience of heritage and museums in general (Smith & Campbell, 2015). However, it is not until recently that the interest in the role of emotions and affective practices has started to grow within museum and heritage studies (Smith et al., 2018). As an example, Andrea Witcomb (2014, 2015) suggests a "pedagogy of feeling" for museums in response to Tony Bennett's (1995) notion of a "pedagogy of walking". In contrast to traditional forms of pedagogy, Witcomb brings forth "nonrational forms of knowledge, ones based on other bodily sensations and on emotional forms of intelligence" (2014, p. 58).

In the field of HCI, the research on affective engagement in museums is growing. Examples of such projects are "See Me, Feel Me, Touch Me, Hear Me" in which mobile technology is used to emotionally enhance a visit to a sculpture garden (Fosh, Benford, Reeves, Koleva, & Brundell, 2013) and "Affective Presence" in which ambient displays in museums are used to augment experiences of affective presence (Boehner, Sengers, & Gay, 2006).

3 Approach

This work follows the methodology of performance-led research in the wild (Benford and Giannachi 2011), in which professional artists and designers directly engage museums and their audiences to create and deploy visiting experiences that are then studied and documented by researchers. The experience at the center of the study - *Sensitive Pictures* - was created by the Serbian artist/designer Bogdan Spanjevic and his company NextGame, in collaboration with the Munch Museum as well as researchers from the University of Nottingham, Uppsala University and The IT University of Copenhagen.

The design process was initiated following the periodic review from period 1, in which the commission asked for a more thorough exploration of emotion detection technology. On 21-22 May 2018 a workshop was organised by the Mixed Reality Lab at the University of Nottingham, in which Spanjevic participated along with researchers from the IT University of Copenhagen and members of the computer vision team from the University of Nottingham. In this workshop the project participants were introduced to technology from the ARIA-VALUSPA project (https://aria-agent.eu/) as well as related emotion detection technologies developed by Prof. Michel Valstar and his team. It was agreed that NextGame would work to explore design opportunities using these technologies.

At the combined consortium meeting and ARM workshop in Copenhagen 29-31 May 2018, the representatives of NextGame met the ARM participants from the Munch Museum, and agreed to explore a collaboration. Several representatives from the museum travelled to Belgrade to participate in a workshop organised by NextGame in August 2018, in which they used theatre-based techniques to develop ideas for the experience. On 3 Oct NextGame met with museum representatives and researchers from Univ of Nottingham and ITU at the Munch Museum in Oslo, in order to test the technology in situ and further develop the concept. On 26-27 Nov a prototype of the concept was tested on ARM participants and other researchers from the consortium.

As a result of the cross-consortium collaboration in these tests, the team developed the idea for Emotion Mapper, which was implemented during spring and summer 2019. On 1-5 July 2019 NextGame and ITU met with researchers at the MRL to evaluate the state of both Emotion Mapper and the Sensitive Pictures prototype (D3.3) and plan a deployment in the Munch museum. An integrated prototype was created over the course of the summer, using Emotion Mapper together with a web app created with the MuseIS platform and the Panopticon emotion detection platform. The prototype was tested onsite at the Munch Museum 1-2 August 2019 with museum representatives as well as a number of museum visitors.

Following some further development work, the final prototype was deployed in the museum 28-31 August 2019. Researchers from all the university partners participated in organising the final deployment.

4 Sensitive Pictures

The design of the Sensitive Pictures experience is presented in deliverable D3.3. In the following, we present some of the main elements as they were presented to users in the deployment in the Munch Museum on 28-31 August 2019.

4.1 Experience design

Six of Edvard Munch's most famous paintings were selected to become part of the intervention: The Scream, Vampire, Madonna, Self-Portrait with Brushes, Christian Munch on the Couch and The Sick Child. Upon entering the museum, visitors encountered a table laden with headphones and postcards displaying these six paintings. Team members wearing custom t-shirts invited visitors to try out the "Sensitive Pictures" experience. Those who showed an interest in participating in the study were briefed by members of the research team, given an identifier card, assisted in opening the web app on their phones and given a small introduction to the interface. The visitors were handed study information sheets and informed what kinds of data the system stored about them, and that they would be given the opportunity at the end of the experience to decide whether to "donate" their data to our research, or to have all data deleted.

The identifier card provided to each unique user contained a three digit code required to begin the experience, and the user was prompted to hold onto it throughout the duration of their experience. If the user had arrived as part of a group each member of the group was encouraged to participate individually, using their own devices. Those who did not have their own headphones were offered to borrow a set from us. Once set up, the user was able to freely navigate the museum in their own time, following their own trajectories.

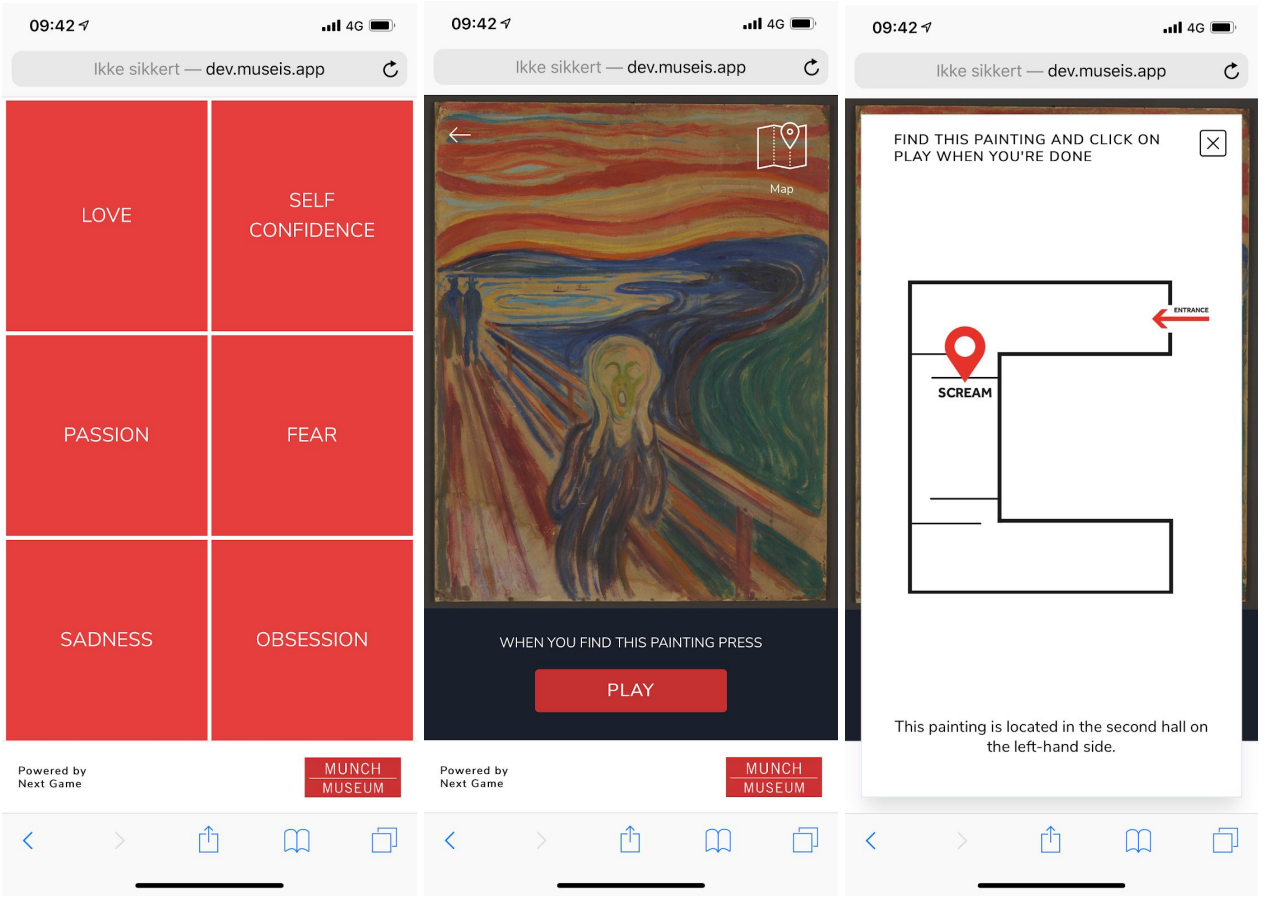


Figure 1: Screenshots from the Sensitive Pictures web app.

The initial screen of the Sensitive Pictures web app offers users a choice of six emotions, laid out in a grid: “love”, “self confidence”, “passion”, “fear”, “sadness” and “obsession” (Figure 1). When a visitor chooses one of the six emotions - in this example, “fear” - the screen displays the corresponding painting - *Scream* - and instructs the visitor to locate the painting in the museum. Visitors may also see the location of the painting on a map. Once the visitor has found the painting and pressed ‘play’, they will hear an audio file which presents a fictional story connected with the painting.

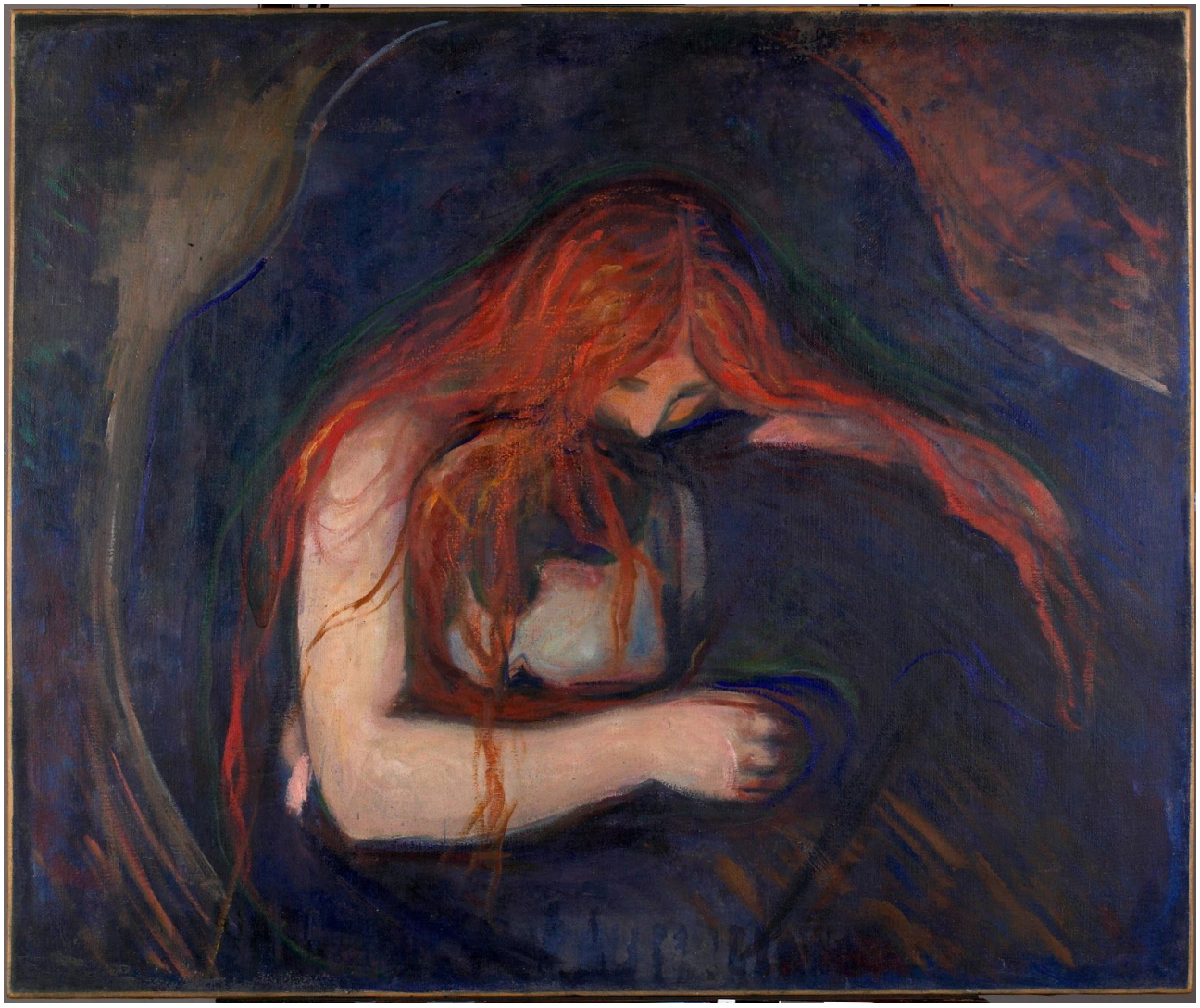


Figure 2: "Vampire" by Edvard Munch. (Photo: CC BY 4.0 The Munch Museum)

For instance, for the painting *Vampire* (Figure 2), the story starts as a dialogue between a woman and a man, with soft piano music in the background:

Woman: He doesn't pay enough attention to me.

Man: I've been devoted to her for years!

Woman: He doesn't love me.

Man: I gave her all my love.

Woman: He is never at home! I miss him.

Man: She'll drain all the life from me! As if she is going to eat me.

> Woman: Look at my arms, hugging this man. It is love!
>
> Man: It's pain. Look at my neck!
>
> Woman: He can hide in my hair. I will kiss him, and all his troubles will be gone.
>
> Man: That's not hair. It's blood! And not a kiss – but a bite.

After the end of the dialogue, the music changes and a new voice addresses the listener in a factual tone, similar to an ordinary audio guide:

> Vampire was painted in 1893, and is considered today as one of Edvard Munch's most renowned motifs. The picture has not always been titled Vampire. Initially Munch gave it the name *Love and Pain*, and it was actually one of Munch's good friends in Berlin, the Polish writer Stanislav Przybyszewski, who suggested *Vampire* as a better title.
>
> Now think about the most intense relationship you have been in. Describe how you feel about it: Whether it is a pleasant or an unpleasant emotion, and how strong is that feeling.

At this point, the screen displays a text input field as well as two sliders, prompting the user to describe their feelings with words and by adjusting the sliders. These sliders represent the dimensions of valence and arousal, following the "affective slider" model of Betella & Verschure (2016).

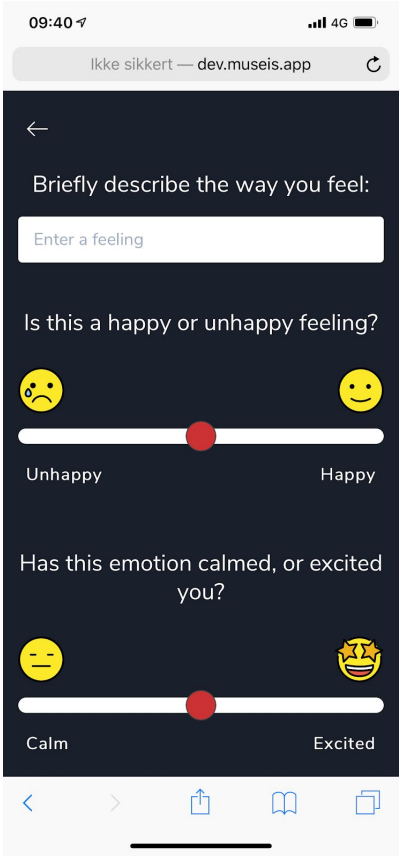


Figure 3: Interface for self-reporting emotional responses.

Once the user had finished using the web app, they were asked to return to the table by the entrance. There they were invited to step behind a divider screen and take a seat in front of a large monitor mounted on a painter's easel, next to a table with an old fashioned telephone (Figure 3). They were instructed to insert their identifier card into a small box under the screen, at which point the telephone would start ringing. Once the user answered the phone, the screen lit up and displayed one of Munch's self-portraits, while a voice on the phone greeted the visitor and presented himself as Edvard Munch.

The voice on the phone would talk to the user about their experience and emotions. Drawing on data from the web app, the experience would speak specifically about the painting that the visitor had reported the strongest emotional reaction to (as measured by the slider representing arousal in the self-reporting interface), and the script was adapted to whether or not the visitor's reaction to that painting was positive or negative (as measured by the slider representing valence). Munch would also ask the user questions about their experience, and the user was given time to respond aloud. During the conversation, a camera mounted above the monitor would register the visitor's

micro-expressions using software from the Panopticon project (for information about the technical setup, see the following section).

Figure 4: The phone conversation with Edvard Munch.

Following the completion of the phone call with Munch, the visitor returned to the desk at the beginning of their visit. Once they handed their card over to the researchers, a souvenir postcard was printed for the visitor. On one side, the postcard showed the Munch painting the visitor had reacted most strongly too. On the other side, there was a visualisation of their slider data, along with the descriptive words they had given at each of the six paintings.

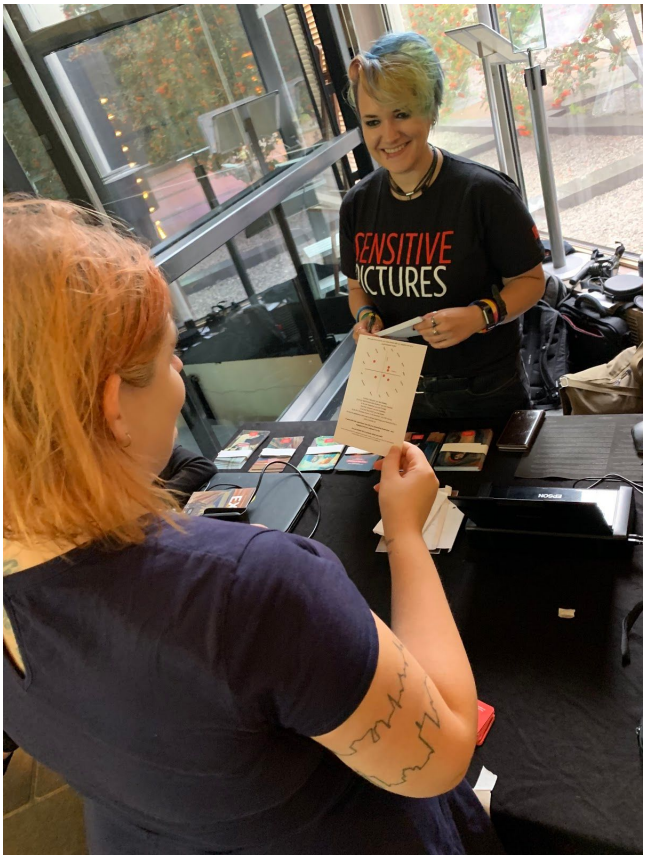


Figure 5: Receiving the postcard at the end of the experience.

4.2 Technical implementation

The setup of the Sensitive Picture experience included three technical parts: MuseIS, Emotion Mapper and Panopticon.

MuseIS

The first stage of the Sensitive Pictures experience was MuseIS, which is a web application that was created and hosted by NextGame. This is described in deliverable 3.3.

Emotion Mapper

The Emotion Mapper component provided the data storage and final stage of the Sensitive pictures experience. The design and development of Emotion Mapper has been presented in deliverable 6.4.

Participant data that was gathered using the MuseIS, and Panopticon components was stored on a UoN server. Using a modified version of Emotion Mapper's admin interface, the component

retrieves data from the other components using various API calls, then stores the data onto the Emotion Mapper database, where the Sensitive Pictures instance of Emotion Mapper is hosted.

Emotion Mapper also created the individual visualisations that were presented to the participants at the end of the experience. The visualisations were created from taking the answers the participants input during the first stage of the experience, MuseIS, along with the results from the second stage, Panopticon, and printing them on a postcard for the participants to keep. Below is an example of one of the printed postcards:

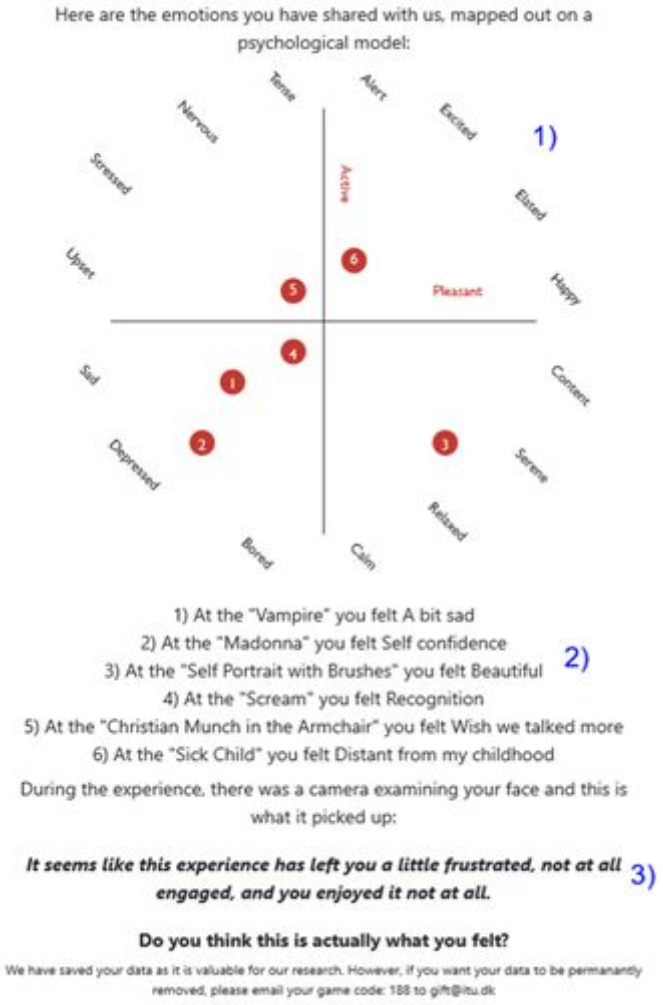


Figure 6: Postcard with visualisation of emotion data from one user.

The first thing that is displayed on the postcard is the visualisation that is generated by Emotion Mapper from the first part of the Sensitive Pictures experience **(1)**. It mapped the Valence (Positive/Negative), and Arousal (Active/Inactive) values that were gathered. These values where then combined to place the painting onto the Circumplex model of Emotion, showing what emotion corresponded to the values the participant self-reported.

Next, the participants see the key for the various points that have been plotted onto the Circumplex model, where each line contains a number and a sentence **(2)**. The number dictates the order in which the painting was visited, whilst the sentence is the free text the participant input during the first part of the Sensitive Pictures experience. This allowed the participants to compare the values they input for Valence and Arousal with the free text input they wrote.

Finally, participants would see the interpretation of what the Panopticon module of Emotion Mapper picked up from them **(3)**. The interpretation used the three values discussed below to form a way for participants to understand what was gathered from them.

For more information about Emotion Mapper, see deliverable D6.4.

Panopticon

The second stage of the Sensitive Pictures Experience was inspired by the UoN Panopticon Ludis project[1] which employed computer vision (CV) techniques to identify and measure the social signals exhibited by museum visitors, specifically their level of Engagement, Enjoyment and Frustration. Specifically, Panopticon Ludis combined CV techniques with a privacy focused experience design, which aimed to make visitors aware of the data capture, equip them with an understanding of the value and importance of their own data, and make them the final owners and arbitrators of how it was used. By using a physical token which was the key to the data, the visitors could explicitly opt-in and out of the data collection on a per-exhibit basis, and upon leaving the venue could make the final decision to destroy or contribute the data to the venue.

The Sensitive Pictures Experience adopted part of this design, by adapting the CV social signal capture for the Munch venue in conjunction with the Emotion Mapper capture. The experience visitors were each given a ticket in the form of a card which was unique to each visitor and had an embedded NFC tag. Following their interaction with the MuseIS web app, the visitors where shown to a booth containing a classic candlestick telephone, a box with a slot, a camera, a mini-pc and a large television screen. The telephone was modified by replacing its internals with an Arduino microcontroller, while the box contained an Arduino based NFC tag reader aligned with the slot. This allowed the phone and the NFC reader to be driven by the Unity3D-based control software on the mini-computer, which also ran the camera and the CV software.

Instructions on the screen, and inscribed on the wooden box, encouraged the visitor to insert their ticket into the slot. As soon as they did this the phone would ring. Once they answered the phone a video would play on the screen which would simulate a conversation between the visitor and Edvard Munch. The video would be chosen from a selection of pre-made videos, based on the profile that Emotion Mapper had built from the earlier segment of the experience. During the conversation the CV software, Bluemax, would measure the Engagement, Enjoyment and Frustration of the visitor and would add it to their profile. As soon as the conversation was over the screen prompted the visitor to retrieve their ticket from the box and return to the onboarding station.

1. https://www.horizon.ac.uk/panopticon-captures-and-examines-video-data/

5 Test results - August 2019

Sensitive Pictures was offered to visitors to the Munch Museum from 28-31 August 2019. 132 visitors completed the entire Sensitive Pictures experience, while a further 65 engaged with part of the experience. On average, visitors spent 46 minutes using the web app to view and respond to five paintings.

The following types of data were collected during the test:

- Self-reported emotion data from the web app (free text answers and quantitative measures from the self-assessment manikin)
- Computer vision data from Panopticon
- Semi-structered interviews with 37 participants, who were approached after finishing the experience.

Every participant received a "study information sheet" when being introduced to the experience. At the end of the experience participants were asked whether they would be willing to "donate" the data collected by the system, to be used in our research. Out of the 132 who completed the experience, only one declined to donate their data, which were then deleted. For the 65 who participated but did not complete the experience, all data were deleted as they did not reach the end station and thus could not actively consent to donating their data.

Interviewees were asked to sign an additional consent form relating to interview data. Interviews were audio recorded and later transcribed.

The photos used in this report have been staged for illustration purposes. The persons depicted are researchers participating in the project who have consented to this use of the photos.

5.1 The user experience

Overall reception

The main takeaway from the trial is that a majority of the interviewees really enjoyed the experience. They described it as very emotional and touching. Some participants even expressed surprise about just how emotional it became for them. As User 160 explains: "*When you told me to go around look at the paintings, I was expecting more of a historical or intellectual sort of exercise but it became a very emotional, and talking about feelings. I liked it a lot*". And as User 233 says, "*It was very touching for me. I never expected that*." The fact that the audio narrative connected the user's own personal life with that of Munch's, was part of what the participants mostly appreciated. User 115 describes it in this way:

> *I think the most interesting part is maybe that I could connect my own story with Munch's story and that was like a parallel line through the exhibition. I didn't expect the camera to turn on me and ask me these kinds of questions. I was surprised, a little bit shocked, a little bit hesitant and when I look back, I'm not sure how comfortable it was. But I think it was interesting to have this kind of parallel line going on*.

Having to walk around with headphones seems actually to have helped in creating a private space where this type of experience could take place. As User 233 puts it, *"I could close myself in. That was astonishing because normally with so many people around I can't concentrate very well and I feel kind of going away. But in this particular time, I was very calm and concentrated".* The highly evocative interpretation of Munch's paintings offered new perspectives and opened up for different ways of experiencing the art, which the participants found valuable. User 211 explains, "*I liked it. It asked questions that I wouldn't have considered otherwise. Interpretations of paintings that are different from my own. It's always good to engage with other ideas and see where it takes you*". And as User 214 says, "*Oh, I think it's great because it made me feel like I'm in the same world with Munch*". The interpretation also led to some confusion in a few of the participants. As User 197 puts it:

> *There is a difference between looking at the picture and hearing the content that was recorded. The recorded content is very directive. So, it was in a way a little bit confusing which I should relate to more, the directive speaking content or the picture itself. The verbal content and the visual content were not always in my opinion necessarily connected.*

But overall, the participants welcomed the possibility of experiencing a more personal connection with the exhibition. In the following we present four main insights from the interview data:

1. The six themes represent a large variety of personal experiences
2. Describing emotions are challenging
3. People try their best to make sense of emotion data
4. Skepticism and interest to the idea of computers trying to read emotions

The six themes represent a large variety of personal experiences

In the interviews, participants shared stories of highly personal and emotional experiences. These seem to have been triggered through a combination of listening to the audio narrative, looking at the paintings and receiving questions to reflect upon. Which paintings and narratives that participants had the strongest reaction to varied across the whole selection, something which demonstrates that the different themes used in the experience were able to capture a variety of personal experiences. Here are some examples connected to four of the six different themes:

Christian Munch on the Couch (obsession):

> *"The biggest guiding, I felt with the picture of his father which is a pretty distant picture when you just stand in front of it. It's not my father. It's not a person I know. It's just somebody reading a newspaper. But the way the questions were brought up made it very intense and very personal",* User 178.

Self-Portrait with Brushes (self-confidence):

> *"Actually, the picture where of Munch himself where he was standing in his coat and with his brushes, self-portrait that says which emotion do you feel looking at that and talks about how he was so confident and how those things. And I felt jealous because I would have liked that kind of confidence, years ago",* User 197.

The Scream (fear):

> *"The Scream was really interesting because I have anxiety like as an actual illness and when I was listening to the narrative it made me very nervous and I didn't like the picture much before because it was always scary. But with the information about him saying like it's part of him and stuff, I've come to that conclusion myself as well. So, I liked hearing that about him",* User 165.

Vampire (love):

> *"I'm in a relationship right now and I'm thinking about it a lot because it's kind of difficult at the moment. And then I saw the Vampire, and actually I think that this part of the app was the best because I really liked these two voices, the man and a woman speaking. It was about the difficulties of relationship and the two different sides that are in a relationship. It can be bad, it can be good, and you have to choose. The typical thoughts you have when you're in a relationship when it's not like super easy anymore. So, that's all the emotions that came back when I was staring at the painting "*, User 114.

Describing emotions are challenging

After having listened to one of the audio narratives in front of a painting the participants were prompted by the app to describe their emotions, in the self-reporting interface (see the presentation of design, ch. 4.1). Most of the users would take to time to do this but the reactions to the form varied. Often participants felt that it was hard to put their emotion into just a few words. As User 114 explains,

> "*I think it was very difficult because you could just type in one emotion or like one feeling. And I was standing there like, I don't know, roughly 2-3 minutes to think about one emotion that captures all the feelings I just felt staring at the picture. And I had a lot of feeling so I was like, I don't know what to type in. What summarize all of that? So, it was quite difficult for me*".

To others this challenge was actually enjoyable. User 233 puts it this way: "*Normally I need more time to do that, to bring it to one word or two or three words. It's a little bit difficult, but it was also*

a good experience because I didn't expect me to react so fast. I felt proud of myself". When it came to the sliders it was clear that they caused some confusion as well as frustration. As User 165 says, "*It was hard sometimes to just have like binaries for the emotions, I think. To say if it either calms you or excites you. Especially because exciting... like they both seem really positive, I think, excitement and calmness"*. And User 112 explains,

> "*There was a bar, calm and exciting, wasn't it? Does it feel calm or exciting? I don't know if that fits perfectly actually, because I was always like I'm excited about this feeling. Actually, I was never calm. When I'm really thinking about the way I'm feeling at the moment that's already a reason why I'm excited*".

The pictures representing Calm and Excited seem to have been problematic for some users too. As User 197 puts it, "*There were sliders between calm and excited, right? And the excited was sort of a happy face with sunglasses and all of that. Excited for Munch was not that kind. Excited would be more like something intense happening, not this happy face with the sunglasses*".

People try their best to make sense of emotion data

After the experience was over the participants were given a postcard in which the data captured by the system was presented back to them, both mapping their responses onto a circumplex model of emotions, as well as presenting the emotion detection data in the form of a sentence suggesting their degree of engagement, enjoyment and frustration during the final video experience (see chapter 4.2). Interestingly, participants tended to respond with interest and openness to these data, often quite willing to offer interpretations and doing their best to make sense of the data presented to them. As User 211 says: "*I don't know about frustrated. Yeah, I guess frustrated in the way that you have to grapple with emotions as they come up and sometimes you don't want to*". And User 165 even blames her back: "*I was very engaged. Frustrated was probably more... I have back pain so maybe it caught that. I didn't feel frustrated. I think*". In general participants seemed to accept what they were presented with as true even though it was unclear to them how the computer would know these things.

Skepticism and interest to the idea of computers trying to read emotions

In the interviews, participants were presented with the question whether it makes sense for technology to try to understand our feelings. The responses to this question show some skepticism towards it being possible for computers to understand emotions, along some interest in what this development could lead to in the future. User 110 puts it this way:

> "*I don't know if it's possible because I think there are boundaries for computers. So, if I decide I don't want to tell my emotions or I don't want anybody to know about it and have like a boundary myself. Then a computer cannot cross that boundary. But I think it's still very interesting to see where those boundaries are*."

Some users are more open to the idea that the computer may see something that might be hidden for humans:

> *"So I do think there's a lot to be learned by facial recognition and muscles around your face and your eyes... you know, discerning what people really... what they're feeling inside. We do it subconsciously and unconsciously and so it's very interesting to see that a machine, that a computer can take those muscles and distinguish feelings. That's pretty cool!"* (#160)

Some users also suggest that this technology might be of help to psychologists: "If it is possible to get the computer to understand when somebody is sad but acting happy then it's a huge thing for psychologists."

5.2 Quantitative analysis of interview data

In order to shed some further light on the interviewee's responses, a researcher with no prior involvement with the Sensitive Pictures project carried out a quantitative content analysis (Krippendorff 2019) on the interview transcripts. In figure 7 below, responses to questions in the interviews are presented as agreement with key statements about the experience.

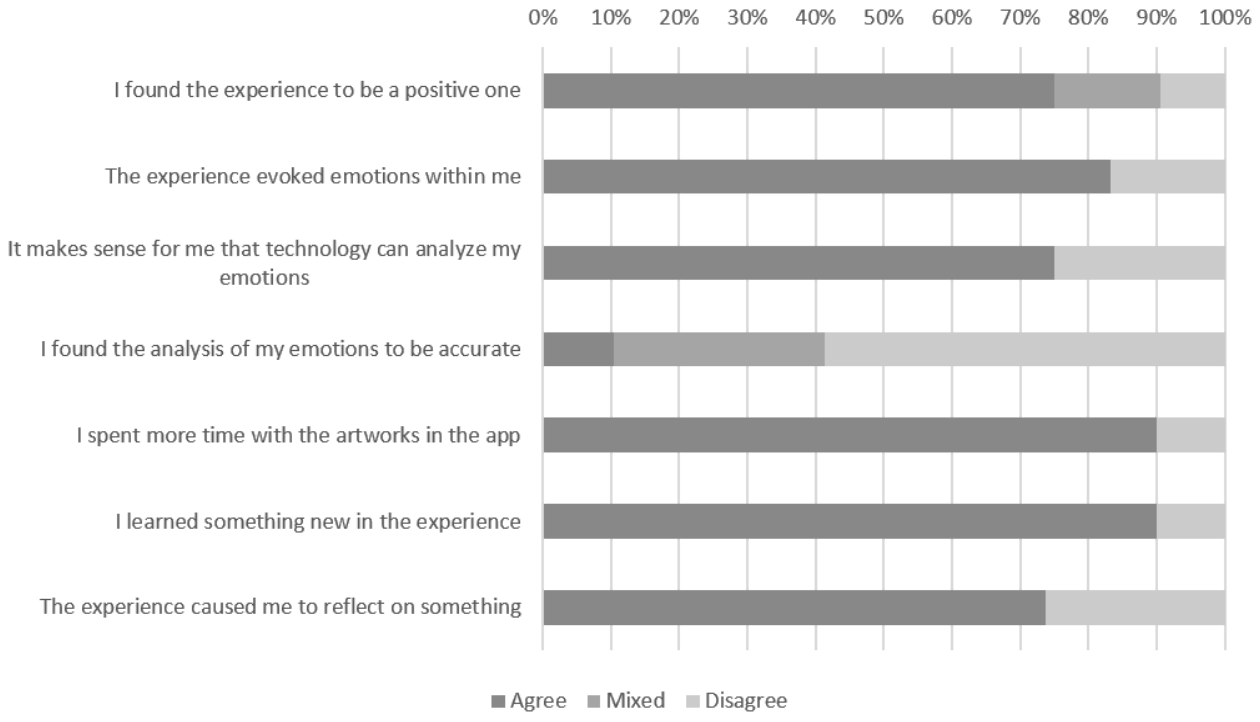


Figure 7: Summary of responses from participants in exit interviews.

Based on the exit interviews we can conclude that most of the respondents found Sensitive Pictures to be a positive experience. Few of the interviewees described it as a mixed or negative experience. Some of these respondents suggested that the audio and storytelling of the app distracted from the immediate experience of the artwork by directing the attention of the participant in another direction.

> *"If it's questions and it's provoking the person to think for themselves and that's fantastic but then with the music. I think it's the painting that should be dominating how you're feeling and what you're thinking, whereas the music is also telling you something else. So you're kind of, all the conclusions and analysis that you draw from this will be like 50 percent painting and 50 percent of what you've, you know, of what you've been told in the art, the music and the excerpt." (#55)*

On the other hand, several other respondents appreciated this interpretative element and multimodal nature of the experience. It is noteworthy that many of the respondents with a mixed or negative impression were museum professionals. As the experience had been designed with ordinary visitors in mind, it is perhaps not surprising that museum professionals - who would likely already be quite familiar with Munch's work - would find it less appealing. As stated by this respondent:

> *"I think that for me that was a massive flaw, because you've told me how to feel before experiencing the painting for myself. And I think anyone that's really involved in art, or are art critics, if they were to use this app, they would not be completely satisfied with being told how to feel before experiencing it for yourself [...] And the best thing about art is that is so subjective. But you've went against what art is all about, it being so subjective, and you've told people how to feel before experiencing the painting" (#55).*

However, not all respondents disliked this aspect of the experience:

> *"In some ways it destroys your own initial reactions because you're listening to someone else's reactions before you even have a chance to think about what yours are. But I liked it because it was a different perspective. Yet I wasn't being lectured at, I was being taken into another world. So I liked that" (#81).*

> *"To start with, I thought, it might be a bit distracting from, like your own experiences, because it's sort of telling you things rather than feeling your own feelings. But you can sort of, you can take it off and have your time as well" (#86).*

Others considered the scaffolding a distinctly positive aspect because it can

> *"give the people a new kind of experience in a museum. But also to ask even some questions. I can understand that you think that most of the people inside the museum, maybe they don't have that experience that they should have. So that you are helping them with to to take you a new direction a different direction from the usual one" (#89).*

The respondents described the experience of Sensitive Pictures in diverse ways. In table 1 below we have identified different descriptors and have grouped them into six categories - four positive categories and two negative categories:

Table 1: Categorization of descriptors used by respondents to describe their experience with Sensitive Pictures. Number of instances of each descriptor is presented in a separate column on the right.

A		B		C		D		E		F	
Good	5	Relaxing	1	Enjoyable	5	Interesting	11	Confronting	2	Frustrating	4
Nice	6	Calm	2	Fun	9	Intriguing	1	Too personal	2	Confusing	4
Great/ amazing	6	Serene	1	Disneyesque	1	Inspiring	1	Like an exam	1	Irritating	1
Impressive/ crazy	2	Warm	1			Surprising	4	Directed	6	Boring	1
Intense	3	Touching	1			Unique/ special	2	Artificial	1	Painful	1
								Political	2	Silly	1
								Intimidating	1	Strange/ awkward	7

The descriptors for the experience ranges from the idiosyncratic, such as 'Disneyesque' and 'like an exam', to more common ones such as 'fun' or 'directed'. Category A groups positive descriptors of a more general phatic nature. Next, category B groups descriptors of a more embodied nature. Apparently, such impressions are less prevalent, or at least, articulated less. The experience of art exhibitions is very subjective. The characterization of the experience as 'calm', for example, stand in stark contrast to another respondent that found it all "*a little bit hard to relate to, because there were several things that were moving, but not exciting like that. In that sense. I don't think anything about his paintings is calming. Any of them*" (#197).

Category C groups the descriptors that characterize the entertaining aspects of the Sensitive Pictures experience and finally category D groups descriptors of a more intellectual stimulating nature. Apparently the experience addresses both the intellect of visitors, but also the more immediate need for entertainment and fun. Several of the participants were surprised as the experience turned out to be something completely different, than they expected.

> *"Like suddenly I didn't expect the camera to turn on me and ask me these kinds of questions. I was expecting to look at the painting, what is that, you know, what is the atmosphere of it? What does it make me feel? But the question was a lot more wonder. 180 degrees other than I expected. I expected maybe 45 degrees. A small link between you and the painting, but not, you know, just like turnarounds: What is your childhood? Or something like that, you know" (#115).*

This respondent clearly expected the more passive experience of a traditional audio guide and was surprised when suddenly the experience turned out to be interactive and she should contribute in a personal way. This respondent was pleasantly surprised, however for others the questioning left a negative impression. Several of the descriptors in category E address this aspect as being 'confronting', 'too personal' and even 'intimidating'. Another important aspect of this category is the impression that the experience was 'directed', preventing the visitor from having spontaneous experiences as discussed above.

The other negative category, F, groups descriptors addressing different dimensions of annoyance, where the experience is the source of frustration and confusion. The one respondent characterizing the experience as painful was because of the high volume on the sound in the concluding conversation with Munch. This last part of the experience sparked the most divergent experiences as some found the situation 'strange/awkward' or 'silly', while others were elated by the opportunity to "speak" to the dead artist: *"Very special, because Munch he died maybe 50-60 years ago, but the video makes me meet him. So it was very special*" (#103). Others were less impressed with the video technology:

> "*Yeah, I think my expectation was that I would meet Munch, kind of, like, yeah you know. I don't know why? I had this impression that he should be a little bit more alive*" (#115).

Overwhelmingly, nearly everyone report that they learned something new during the experience and most also report that the experience made them reflect something they hadn't thought about before. The learning enabled by the app also corresponds with nearly all respondents reporting that they spent more time with the paintings featured in the app, than with artworks from the rest of the exhibition.

Emotions

Nearly all the respondents agreed that the experience evoked personal emotions for them. The awareness that someone, even a computer, is paying attention to your emotions make you focus on them as well and perhaps even influence them. As stated by one respondent:

> *"So I think I was aware that I'm maybe being watched or that my emotions count. So you think more about them. So as soon as you know that somebody is paying attention to your emotions you pay attention to your emotions as well. And I think you can sometimes, maybe, change your emotions a bit from what they would be if you didn't know that somebody was interested in them."* (#114)

Respondents rarely specify in the interviews which emotions are evoked. Therefore, the range and frequency of emotions following from the experience is surely more comprehensive than what is presented in table 2 below, where we have identified 22 different emotions specified by the respondents.

Table 2: The different emotions mentioned by respondents.

Positive emotions	Negative emotions
Happiness	Guilt
In control	Unhappiness
Empathy	Unsettled
Love	Dread
Nostalgia	Caught
Calmness	Fear
	Despair
	Sadness
	Anger
	Embarrassed
	Shock
	Hesitant
	Loss
	Put on the spot
	Anxious
	Lost

The emotions specified by respondents are predominantly of a negative, or difficult, nature. The quoted statement above indicated that there might be an element of social desirability bias, when you know, your emotions are being monitored, which could entail hiding or suppressing difficult emotions. But this does not seem to have been the case with regard to reporting on the experience. However, several respondents found it difficult to put their emotions into words:

> *"And I think it's very difficult, because you couldn't just type in one emotion or like one feeling, and I was standing there like - I don't know - roughly 2-3 minutes to think about one emotion that captures all the feelings I just felt staring at the picture and I have a lot of*

feelings, so I was like, I don't know what to type in! And I was like, okay what summarizes all of that? So it was quite difficult for me." (#114)

Others relished the opportunity to be specific:

"I think it was great that the first box you could put in the emotions that you felt and so you could be very specific, because you could write the words yourself. So, for instance, the first painting that kind of provoked, like, perhaps some sadness, I could be more specific and you could say 'sorrow' or 'solemn' or 'subdued'. And so I put down very specific words, so that you could really interpret and make a proper analysis about how I felt." (#55)

More importantly, even though the experience evoked difficult emotions, most respondents reported it to be a positive experience. Apparently, the experience offers visitors an opportunity to experience difficult emotions, without being overwhelmed by them, and perhaps even reflecting on them as most respondents reported that the experience was a cause for reflection. Some respondents felt the experience had a therapeutic aspect: "*It was a strong emotional experience. In fact difficult to answer questions, because it's just kind of a blurb of emotions. Hard to really put into words, but it's kind of a storing box... And kind of, a little bit like a therapy.*" (#80) "*I thought the narration was really, sort of like a psychologist putting you into trance.*" (#70)

5.3 Personal data and trust

Trust within galleries and museums is a valuable commodity, one that fosters long term relationships, increases the volume of visitors, and encourages deeper engagement (Passebois and Aurier, 2004). Trust in museums is also important, due to their position in society as spaces which are "constantly changing and complex political entities shaped by the society in which they are situated, including the perspectives of their visitors" (Knell et al, 2007, cited in Lynch, 2013: 6). This means that galleries and museums have power of representation regarding the world around them, how people perceive that world, and how people frame their discourse around their perceptions (Dodd and Sandell, 2001). This gives museums a prerogative to constantly reflect on their exhibits and adapt to the requirements of modern life (Duncan, 2002), but also a unique freedom to try experimental ways of representing contemporary society.

The Munch museum is a dedicated institution which displays the life's work of Edvard Munch. The name of Edvard Munch carries with it certain perceived qualities; most people who visited the museum had at least heard the name and knew of one or two of Munch's paintings as a minimum, even if they had not visited the museum before. For example Munch's *The Scream* is an iconic and globally recognisable painting which many visitors prioritise during their visit to the museum. Therefore, the Munch museum can be seen as a stable and trust-worthy home for a widely renowned artist. Passebois and Aurier (2004) identify this 'perceived quality' as the first stage in creating trust and commitment between audience and venue. 'Perceived quality' includes quality of the contents being exhibited, quality of interaction and quality of environment, all of which

build the first foundational block of a trust-based relationship. If expectations of these qualities are met or exceeded, then the next stage of the relationship is reached: perceived value. Perceived value explores the relationship between the audience and the venue in terms of hedonic value (or enjoyment); cognitive value (educational potential); spiritual value (escapism); bonding with art, venue and peers; and social discrimination (cultural capital, elitism, self-distinguishing). Once again, if these requirements are met, the next stage is reached, cumulative satisfaction. Cumulative satisfaction is the final stage, according to Passebois and Aurier (2004) before long-term loyalty and trust is established. This accounts for the previous two stages, and represents the enchantment or surprise experienced by the visitor. Alongside with institutional reputation, these stages contribute towards a trust-based relationship between visitor and venue.

Because the museum exclusively houses Munch's works, it already has a high level of perceived quality both in terms of the exhibit, and the physical environment. Perceived value, the next stage in the Passebois and Aurier (2004) framework towards the creation of trust, was the core stage that the Sensitive Pictures app integrated with. By affecting the hedonic, cognitive and spiritual values, as well as encouraging bonding with Munch's work, the app worked to build on the perceived qualities already delivered by the museum. This paved the way for increased cumulative satisfaction, and therefore increased levels of trust. The attitudes towards the app from those who participated in the study demonstrate the success of this process:

> *"for me like coming into this exhibition was much more relatable than any other art exhibition that I've been to and that is a personal thing I suppose. I don't really get a reaction from some flowers that I do from perhaps like a spirit or a story that I may see in the paintings of one of Munch (…) So for me this was a really fantastic experience"*

The ability for Sensitive Pictures to effectively impact perceived value was achieved with the introduction of two core types of technology to the museum environment; facial recognition, and emotion data capture software. 'Emerging' technologies such as facial recognition are often contentious whilst the public determine the societal value of those technologies (Macnaghten et al., 2019). When technology is hidden, not well understood, or feels 'forced' on the public, general narratives surrounding that technology are often perceived to be negative (Macnaghten et al., 2019). This is, to some degree, the status of facial recognition technologies or emotion capturing technologies right now, due to covert practices in public space and social media sites. As one participant said:

> *I think, you know for me today, I heard on the news that Siri on your phone is listening and recording your talk questions and... So Apple sends it to their, you know, collaborators and things like that. So you have this kind of resistance a little bit these days to giving away your story because you never know where it will end up somehow. So it's a little bit like that.*

However, this project consciously removed several of these barriers, establishing trust with the user by clearly explaining what data was being captured, how it was being captured, and how the

user could withdraw from the study. This transparency in the research process broadly allowed the user to engage with Sensitive Pictures with minimal pre-determined social bias, although this was not the case for everyone. Participants were therefore able to judge the reliability of the technology based on the outputs given to them on their postcards. This process was also impacted by the existing authority of the museum itself, and of the team of professional researchers, both of which command certain levels of trust, even if that trust largely represents a duty of care, and a responsibility to uphold ethical and safe practices that will ultimately benefit the user (Nzinga, 2016). This preconceived expectation of care allows museums freedom to experiment with technologies that may not yet be fully integrated into society, with decreased levels of pushback from the user.

During interviews, as part of establishing public attitudes towards these technologies, the participants were asked if, based on their experience, they believed it made sense for technology like facial recognition, or emotion data capture, to try to understand their emotions. Only relatively few participants said that they trusted the technology to reliably capture that information:

> *Yes. I think you can ... It's amazing that you can use technology to assess mental and emotional states. In a like sort of on the spot kind of way*

Several participants said that they thought it was possible for technology to capture some emotions, but not reliably due to the complexity of human emotion:

> *Participant A: I mean like the computer is still programmed by human beings. I think you guys just looked like somebody who was smiling and then like looked at the face and you were like OK if the face does that that's happy so you taught it and like it for computers. I think what it, what I, what do I want to say? Yeah I mean, like, basically a computer knows what...*
>
> *Participant B: It's still programmed by a human's humanity.*
>
> *Participant A: Yeah it's approved by humanity and maybe because in cases when somebody is sad but still is smiling then it's not working. Is it?*

One participant suggested that they thought it would not be possible to capture emotions at all:

> *I think it's very difficult because you couldn't just type in one emotion or like one feeling and I was very like I was standing there like. I don't know roughly 2, 3 minutes to think about one emotion that captures all the feelings I just felt staring at the picture and I have a lot of feeling so I was like I don't know what to type in . Yeah, yeah, exactly, and I was like okay what summarizes, like, all of that. So, it was quite difficult for me.*

And some participants said that they were not sure if the technology was capable of capturing emotion, but they thought it provided an interesting or useful tool in the context of galleries and museums:

...people aren't really taught the tools to kind of reflect on that and being able, being comfortable to kind of like sit with something and like think through it . And so I think just having, like with like the recordings, kind of like guiding people through that and then like asking them those questions in like a safe space to have them actually like engage with the art [...] I think... that is really poignant and could be really impactful.

Several participants also mentioned that they thought it was an important field for research to be carried out in. For some users, the potential future applications of the technology in other fields was particularly important. For these users, some of the frustration or problems with the app were able to be forgiven, because they trusted that the long term implications of the research were valuable:

I think that's fine, because I know a little bit about how computers are like detecting, you know suicide risk, right. Like psychosis earlier than humans and stuff like that. So I think it's an important field.

Participants were then asked to reflect on the results given on their postcards, whether they trusted that the Emotion Mapper results reflected their input, and if the Panopticon results accurately captured their emotional state during the phone conversation with Munch. Contrary to what participants said when asked if they believed the technology was capable of understanding their emotions, the interviews found the majority of participants agreed with the sentence printed on their postcard, suggesting a high level of trust regarding the accuracy of the technology. This was particularly relevant regarding the Panopticon result.

You never can tell. I. Yeah, I think it's recording to this very true. According to this, to the second, this is revealing I think, for me. Deep inside I think it's true. Yes.

Some participants even changed their reflections upon examining their postcards, particularly in reference to the reading regarding frustration. Having previously stated that they were not frustrated, several users attempted to find ways that this might have been true:

I don't know about frustrated . Yeah I guess frustrated in the way that, you have to grapple with emotions as they come up and sometimes you don't want to.

I am not feel frustrated, actually, but I'm really engaged and enjoy a lot. And... I think I'm not, and I'm not frustrated. How to say? I'm not really frustrated, actually. I just feel very complicated after all this is exhibition because, yeah it's very different for me. This is my first time to see a lot of art in the museum... real famous art in museum. Yeah.

This suggests that for some participants, trust in the facial recognition technology to accurately notice expressive cues, outweighed the trust the user had in themselves to passively do the same thing, or even to accurately self-report their emotional state to Emotion Mapper:

Interviewer: And this is our question to you. Do you think this is actually what you felt? Is it true?

Participant: Yes. Somehow frustrated. Yeah, engaged. Yeah I think it is. This looks more familiar to me than this

Interviewer: So that the output from the camera eyes you believe that more than you believe the…

Participant: Yeah, yeah .

Interviewer: Sliders.

Participant: Yeah

A few users even tried to validate Panopticons results by explaining why the computer might have sensed frustration in them, when they did not consider themselves to be frustrated at all:

Interviewer: And it says... you have been somewhat frustrated, very engaged, and you enjoyed it a lot.

Participant: I did. And I was very engaged. Frustrated was probably more... I have back pain so maybe it caught that. I didn't feel frustrated, I think

Interviewer: So, so if you look at all of this, do you think it tells the truth?

Participant: Yeah.

Only three participants interviewed stated that they were not frustrated, and could not work out why Panopticon had said that they were. Importantly, a very small number of participants felt that their emotional responses to the paintings had been unfairly tainted by the experience, or that the requirements of the experience, for example summarising their emotions in one word or short phrase, negatively impacted the way they interacted with the artwork on display.

But there's not the emotion or me. That's the thing. (...) the best thing about art is that is so subjective. But you've went against what art is all about, it being so subjective, and you've told people how to feel before experiencing the painting. I was really disappointed or annoyed by that actually, personally.

Interestingly, these participants' negative experience with Sensitive Pictures did not seem to affect their overall trust in the museum, the perceived quality of the work to them was much more defining:

Yeah, yeah. And I really kind of felt… I think it's just you can take so much from these paintings you know there's just so many contrasts, but the hues in which he uses them. I think there's kind of a lot of, there's a lot of texture, it kind of seems like a lot of surrealism. There's a lot of reality in it as well and there's these sort of straight lines and these are kind of wavy and at could be I think there's just so much telling you from, from these paintings which has been quite amazing. So for me this has just been a nice

experience to, to have to think about my own generation and how this would apply to his paintings today.

This may be a result of increased perceived value via cognitive value and bonding with Munch's work, as well as the 'surprise' of taking part in a novel experience contributing towards their cumulative satisfaction:

And it was actually quite a bit of a pleasure to see to have to be provoked in some ways and to think more. And then you kind of, you're processing what (...) you will say or how you will think and how well you will put that across. (...) But if you if you really process what you're what you're thinking then that and then you're kind of constructing and making a conclusion for the paintings in your mind.

Every participant interviewed demonstrated varying, but consistently positive, levels of perceived value towards the museum and Munch's art after their experience with the Sensitive Pictures application:

I think as much as I kind of love and hate this kind of technology that reads people, I think it is also interesting for people to have a chance to connect with themself (...) So I think it can be an interesting and also valuable tool in helping people to connect to. Even though you're being observed somehow but I think this kind of technology is interesting. Yeah.

Despite the limitations of the deployment, an unexpected outcome of Sensitive Pictures is the reflection on levels of trust. Museums and galleries often have pre-existing levels of trust present in the public, particularly larger or well-established museums, due to their assumed ethical and moral practice. Researchers coming from large or well-known institutions also often enjoy the freedom granted by this assumption. The authority and power of these two industries working together allowed for the deployment of an experience, which may not have been as socially accepted within a different setting, because pre-existing trust offered a platform for the general public to interact with these technologies in a way that may be considered more 'safe'. This then allowed them to engage with the technology more completely, for the most part reserving scepticism or dismissal until they had taken part in the experience, and allowing for a more honest engagement with the self-reporting activity. Participants initially expressed a reticence to believe that the technology employed would be capable of capturing and understanding the emotional journey of the visitor through the gallery. However, once engaged with the technology, particularly the facial recognition technology although not exclusively, it became apparent that the participants placed high levels of trust in the outputs, even if those outputs were false and did not represent the actual data capture. The long-term results of this, seems to show an increase in perceived value surrounding the museum and Munch's art, particularly in terms of cognitive value and bonding, but also hedonic and spiritual value as well, ultimately increasing levels of trust with the museum.

5.4 Analysis of emotion detection data

This chapter describes the analysis of participants' social behaviour while engaged in the "phone conversation with Munch" part of Sensitive Pictures. The images captured by the camera were passed to the Panopticon System which consisted of two subunits: the Blue-Max (BM) subsystem and Social Signal Processing (SSP) subsystem.

The BM unit analyses the facial actions (i.e. facial muscle activations) and head movement activity observed in each image sequence. BM detects sixty-six (66) facial coordinates corresponding to the position of facial components (i.e. eyes, mouth and nose). Using heatmap regression, it estimates facial action unit intensities from these coordinates. Action units in this context refer to the contraction or expansion of specific facial muscles. The heatmap regression framework for action unit intensity estimation was developed as part of the ARIA-VALUSPA EU Horizon 2020 project[1] (Sanchez-Lozano et al. 2018). Specifically, BM outputs five facial action unit intensities on a six-point scale ranging from 0 to 5 and three head movement parameters.

The facial action units predicted include AU6 (Cheek raiser), AU10 (Upper lip raiser), AU12 (Smile), AU14 (Dimpler), and AU17 (Chin raiser). Head movement was measured as head pose angles in the vertical, horizontal and lateral directions corresponding to pitch (head nods), yaw (left-right turns) and roll, respectively. Hence, eight parameters—5 AU intensities and 3 head movements—were computed for each image. Next, the BlueMax output is passed to the SSP subsystem which measures the degree of enjoyment, engagement and frustration as inferred from participants' behaviour during the interactive activity.

To analyse the social signals, the head pose signals from BlueMax are first converted into angular deviations. This is done by first computing the mean head pose angles on valid observations only (i.e. frames with detected faces), and then subtracting the mean from each observed head angle. Next, the AU and angular displacement signals are normalized as follows:

$$d_x = \frac{abs\,(\theta_x)}{\theta_{max}} \qquad [1]$$

$$P_{AUx} = \frac{I_{AUx}}{I_{max}} \qquad [2]$$

where θ_x is the observed angular head deviation in any x direction, θ_{max} is the maximum possible deviation, d_x is the normalized deviation, I_{AUx} is the observed intensity for $AU_x : x \in 6, 10, 12, 15, 17\}$, I_{max} is the maximum possible AU intensity and P_{AUx} is the normalized AU intensity. Thus each observation assumes a value between 0 and 1 which could also be interpreted as a percentage value.

Next, the normalized head pose and AU signals are combined into an Nx8 dimensional signal which is then divided into non-overlapping experience segments, each consisting of 10 temporally aligned images/frames (approximately 0.4 seconds). That is, for an image sequence of 500 frames, we end up with 50 experience segments. A segment is excluded from further analysis if no face

[1] https://aria-agent.eu

was detected in all of its constituent frames. Similarly, all participants having less than ten (10) valid experience segments were excluded as these did not contain information sufficient to determine the participants' behaviour. Further, a representative value is computed for each experience segment as the mean taken across all eight signal values.

The SSP models are crafted based on behavioural associations documented in literature: engagement has been found to correlate with postural movements (e.g. leaning forward) as well as facial muscle activity (Hammal et al. 2015); frustration is commonly associated with movements in the mouth region and increased postural activity (Grafsgaard et al. 2014), while enjoyment is usually associated with expressions of happiness, e.g., smiling (Frank et al. 1993). The SSP models each of the social behaviour of interest are as follows:

$$S_g = mean[(d_p,\ d_y, d_r)\ , (P_{AU6},\ P_{AU10},\ P_{AU12}, P_{AU15},\ P_{AU17})\ ] \quad [3]$$

$$S_j = max(P_{AU6}, P_{AU12}) \quad [4]$$

$$S_f = mean(d_p,\ d_y, P_{AU10}, P_{AU17}) \quad [5]$$

where S_g , S_j , S_f are the computed engagement, enjoyment and frustration levels, respectively, $d_p,\ d_y, d_r$ are the normalized pitch, yaw and roll angular displacement, and P_{AUx} is the normalized intensity for $AU_x :\ x \in \{6, 10, 12, 15, 17\}$.

The social signals for each experience segment are computed using Equations 3 to 5. Then further analysis was conducted on the generated social signals to identify behavioural patterns among participants. To do this, two quantitative metrics, f_{active} and f_{mean} , were generated for each participant's signal as follows:

$$f_{active} = \frac{N_{active}}{N} \quad [6]$$

$$f_{mean} = \frac{\sum_{i=1}^{N} S_i}{N} \quad [7]$$

where N_{active} is the number of experience segments in which the target social signal is active, N is the total number of experience segments in the participant's image sequence, and S_i is the signal value of the i_{th} experience segment. Segments with signal values $S \geq 0.5$ are considered active segments while those below the threshold are considered inactive. This resulted in an Mx2 feature vector F where M is the number of participants and the first and second columns of F correspond to the f_{mean} and f_{active} features, respectively.

To identify participant groupings for each signal, clustering analysis was performed on the derived participants' features in two steps. First, hierarchical clustering was performed to determine the

number of clusters k present in the group, then K-means clustering was used to classify the participants into k groups.

Hierarchical clustering analysis revealed different participant groupings for each social signal: four for engagement, three for frustration and two for enjoyment. Figure 1 shows the K-means classification of participants into the identified groups. The y-axis represents the percentage of active experience segments while the x-axis represents the mean signal value for each participant. Each point on the plot represents a participant.

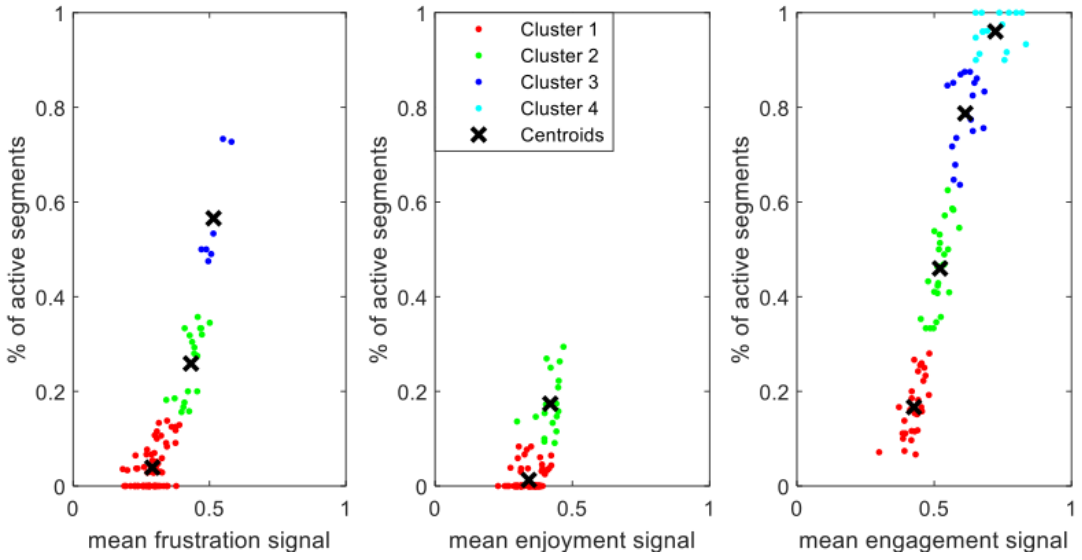


Figure 8: K-means clustering of participants for each social signal based on the observed mean signal and percentage of active experience segments in participants' image data. Each point on the plot represents a participant. 3, 2 and 4 participants' groupings are identified for frustration, enjoyment and engagement respectively.

A careful observation of the plots in Figure 1 shows that the participants appear to be dominantly stratified according to the percentage of active experience segments due to the limited variation in the mean signal values. Since the activity was not very emotive, most of the observed signals fall within the mean region (i.e. neutral), thus reducing the impact of the mean on the clustering analysis. Due to this observation, the user clusters will be defined in relation to the degree to which the participant experienced an emotion or exhibited the target social behaviour during the activity. Definitions of the identify participant groups are presented in Table 1.

Table 3 Description of identified participant groupings based on the assessed behavioural parameters

Social signals	**No. of identified participants' groups**	**Definition of Cluster mappings**

Engagement	4	*Cluster 1*—Mostly disengaged *Cluster 2*—Somewhat engaged *Cluster 3*—Mostly engaged *Cluster 4*—Very engaged
Enjoyment	2	*Cluster 1*—Neutral *Cluster 2*—Enjoyed some bits of the experience
Frustration	3	*Cluster 1*—Did not find the experience frustrating *Cluster 2*—Found some bits of the experience frustrating *Cluster 3*—Found most parts of the experience frustrating

Note that this interpretation assumes a correct measurement of the behaviour primitives (i.e. AU intensity, head movements) although this is not always the case. Also, for most participants, face detection failed in a significant number of images, particularly those recorded on Day 1 and Day 2 which meant that only portions of the experience with valid behavioural information were analysed. Thus, some parts of the interaction with useful behavioural information would have been lost. Therefore, due to the above-stated complexities, it would be necessary to interpret the group definitions with caution as these may differ from the participants' perception of their experience.

5.5 Visualisations from Emotion Mapper

Data from the self-reporting interface in the Sensitive Pictures app were collected on the backend with the Emotion Mapper tool described in deliverable D6.4. In the following, we present analysis of these data.

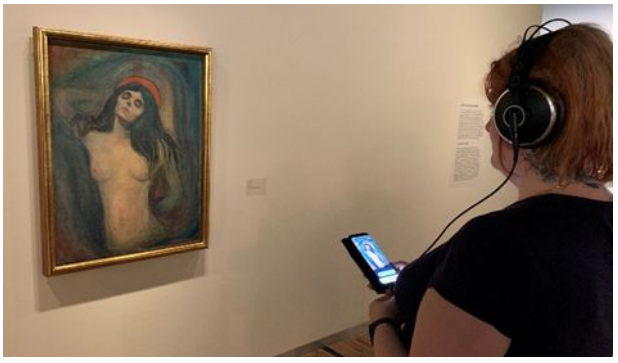


Figure 9: Participant listening to the audio track for the Madonna

The visualisations themselves are interactive and will be made available to the reviewers. The data can not be made publicly available, due to a clause in the consent form stating that the data collected - which potentially might include quite personal information - would not be published.

The following captured images give a quick overview of how the visualisations appear. Figure 9 shows the visualisation control interface with the Sensitive Pictures data loaded. There is a selectable list of all visitors at the top and exhibits (Vampire, Self Portrait of Munch, Madonna, Scream, Christina Munch in Armchair, and Sick Child) down the left. In this example all visitors and exhibits have been selected, resulting in the aggregate emotional summary displayed on the right. The horizontal axis corresponds to valence and the vertical axis to arousal. The text labels around the outside are standard ones associated with the Circumplex model. Each coloured circle shows a different point at which visitors rated their arousal and valence. The circles are discrete because MuseIS chose to employ a discrete eight point rating scale for each axis (though other variants would have been possible too). Positions have been slightly adjusted to make the display more circular. Circle size and colour show the popularity of this particular point on the Circumplex model space. The visualisation reveals a wide range of emotional responses from visitors.

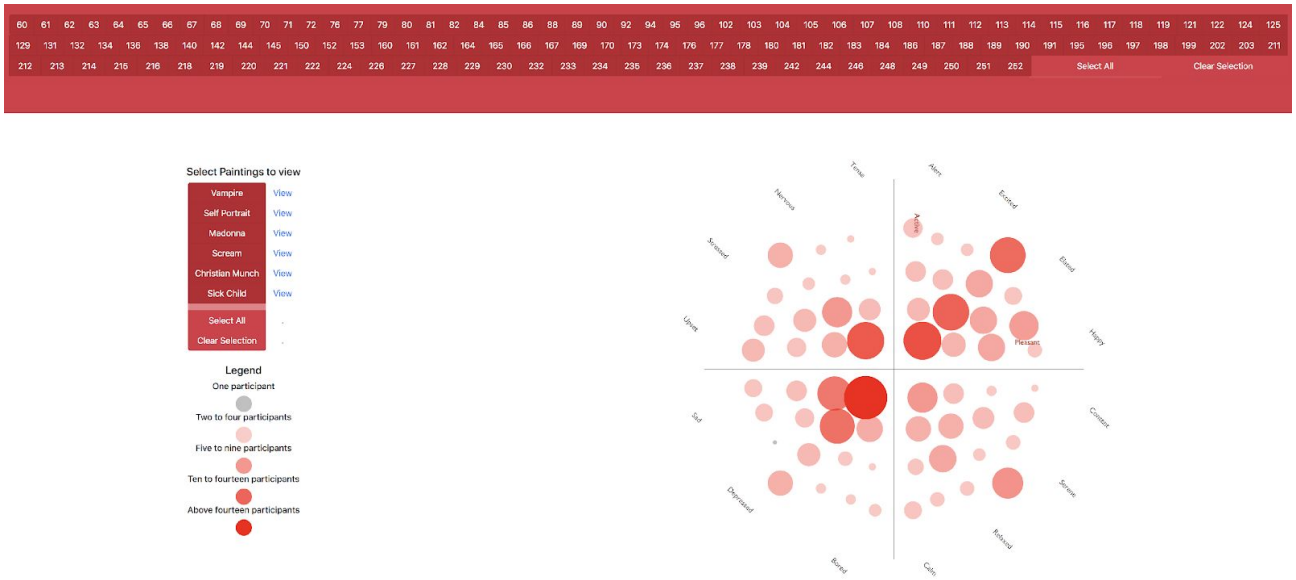
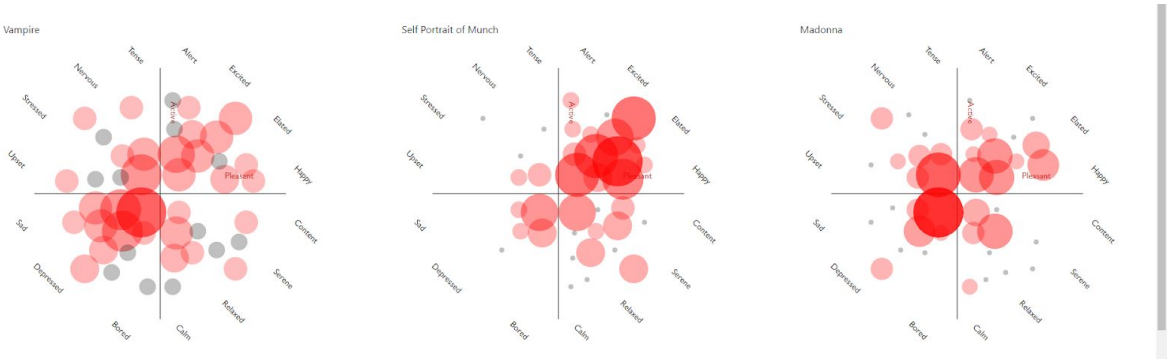


Figure 10: The visualisation control interface.

Users also can mouse over any of the data points to gather more information on the words given in answer to the free text question. The following example shows inspecting a data point that was contributed to by 28 visitors who responded with a variety of words.

Figure 11: Mousing over a data point reveals the number of contributors and the words they used.

The following visualisations show exhibit profiles for the six audio experiences delivered at six Munch paintings (clockwise from top-left: The Vampire, Self Portrait of Munch, Madonna, Sick Child, Christian Munch in Armchair, Scream).



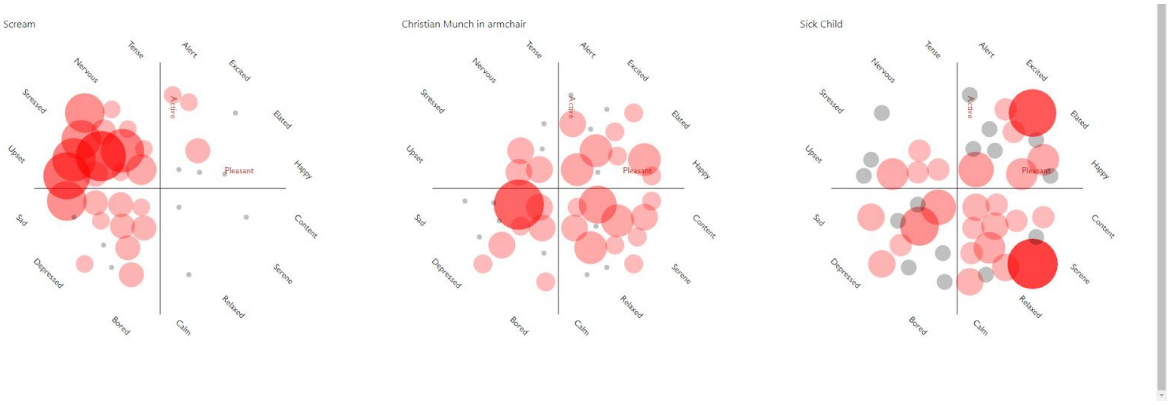


Figure 12: Six exhibit profiles.

These invite interpretation as to whether visitors did tend to react in particular ways. The Scream appears to have an overwhelmingly negative valence, whereas Self Portrait tends to the positive (and aroused perhaps?). We suggest that it might be useful to present such visualisations back to visitors and invite them to reflect on them – as labels on signage for instance?

One can also create profiles of individual visitors in response to all of the exhibits, a selection of which are shown below. Again, we are not claiming that such a technique scientifically profiles visitors into rigorous categories that would, for example, allow for the recommendation or adaptation of visiting experience – though this might be a topic for future work beyond the project. Rather, at this stage of development, we envisage presenting such visualisations back to visitors as a further stimulus to their own self-reflection or storytelling in relation to the museum. An example of how this might be done is via personalised souvenirs such as the use of postcards, as done in the August 2019 test.

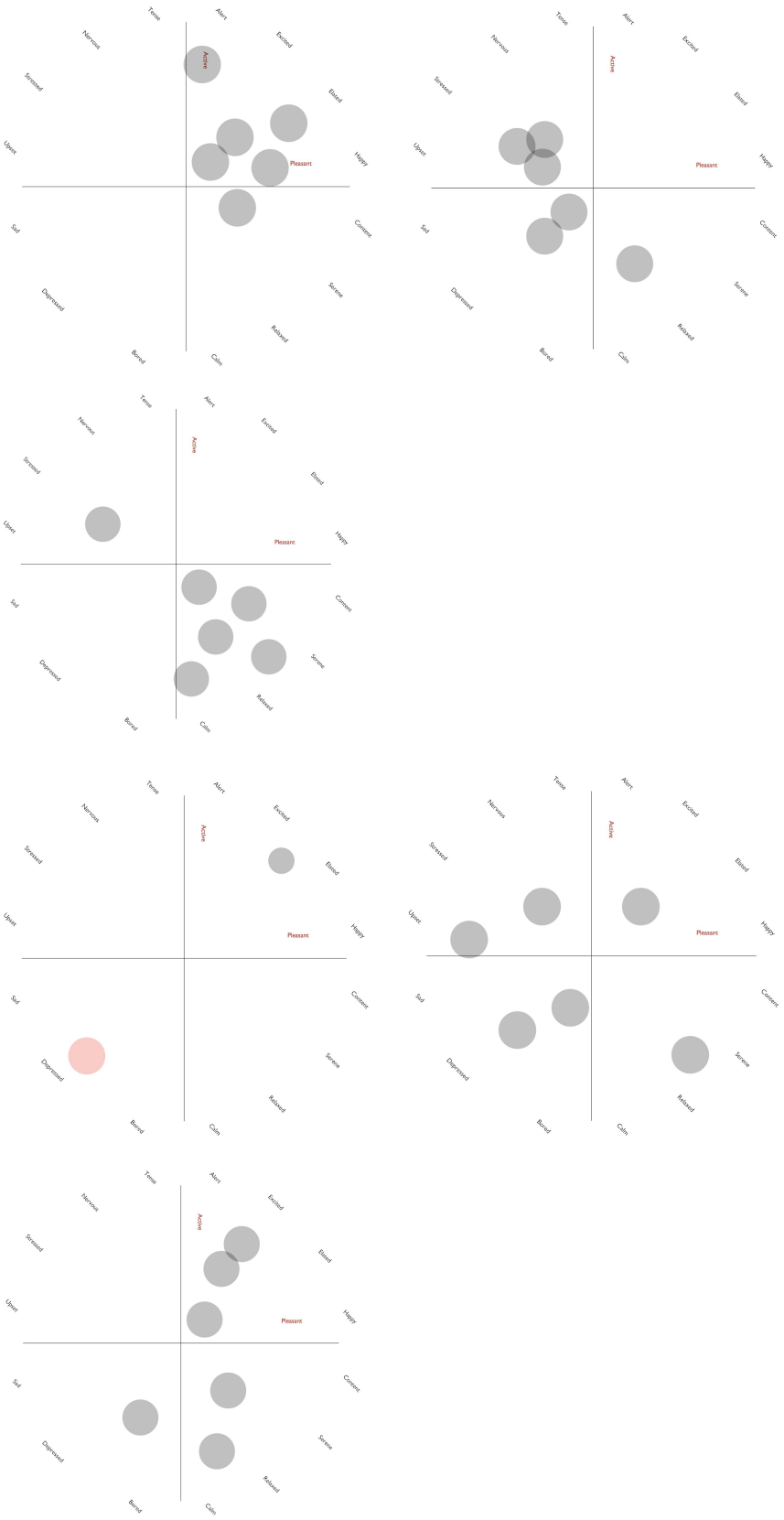


Figure 13: Individual profiles suggest how each visitor experienced the exhibition.

So far, our examples have focused on position exhibits or visitors within an emotional space, in this case that defined by the two dimensions of the Circumplex Model. In addition, we might consider the words that visitors entered in their free text answers to see what emotions these might reveal. One way of approaching this is shown in the prototype visualisation below. This shows a word cloud in which each red circle represents a word that one or more visitors has written. Individual worlds have been separated out and common meaningless ones removed (e.g., pronouns, conjunctives and so forth) to generate a set of distinct meaningful words, each of which is associated with a count (the number of times it has been used across all visitors' responses) as well as average valence and arousal scores (again across all visitors' use of the corresponding sliders). The section top right of the figure shows an overview of the entire world cloud, with each word being represented by a circle, the size of which shows its frequency of occurrence. Dragging and resizing the pink rectangle allows zooming in or out to select a region of this space that is then shown in detail on the right, with each word being represented by a labelled red dot. This representation nudges words that occupy the same position apart a little so that they can be clearly red, but at the cost of a little inaccuracy in positions. Selecting a word then brings up information about it in a text box and also draws its 'emotional footprint' that is the spread of all of its registered positions as a shaded area. This is intended to convey some sense of both its position but also spread or focus within the valence-arousal space of the Circumplex model. In the figure we have selected the word 'sad' which occurs 36 times in the Munch dataset and that has a centre of mass and footprint largely in the bottom-left (passive-negative) quadrant, quite close to where advocates of the Circumplex model suggest that it should be located as an emotion.

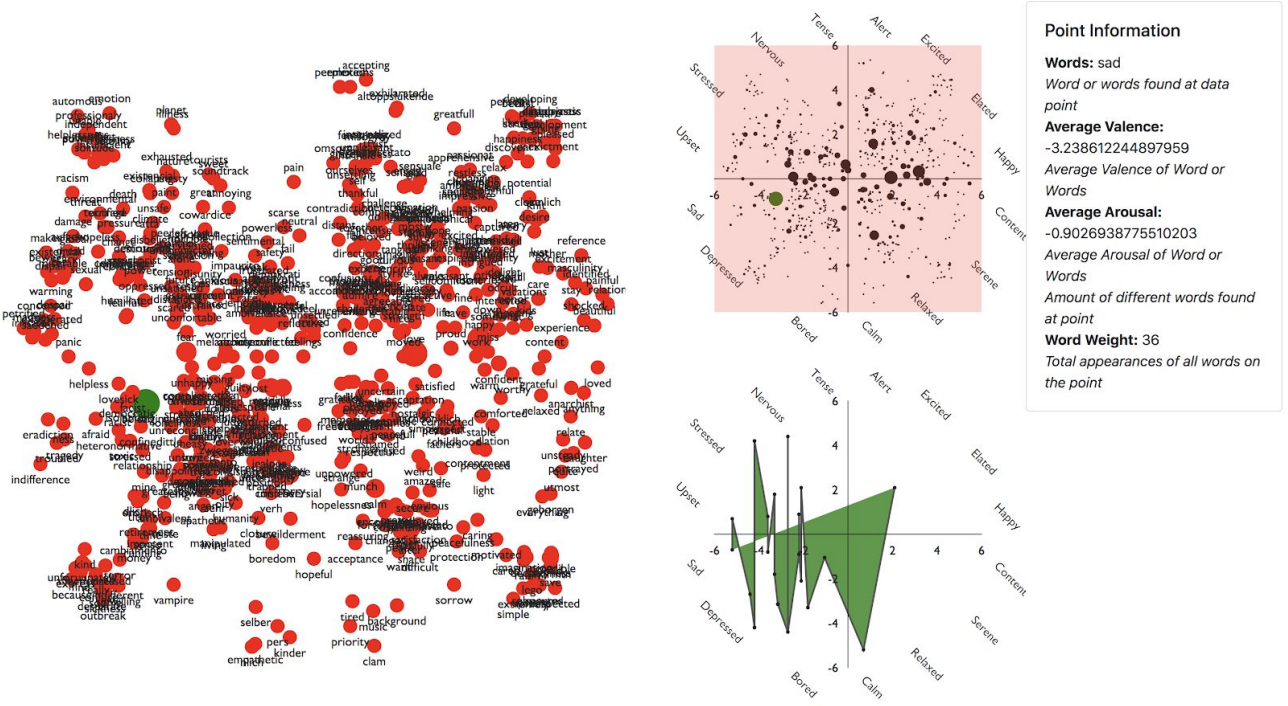
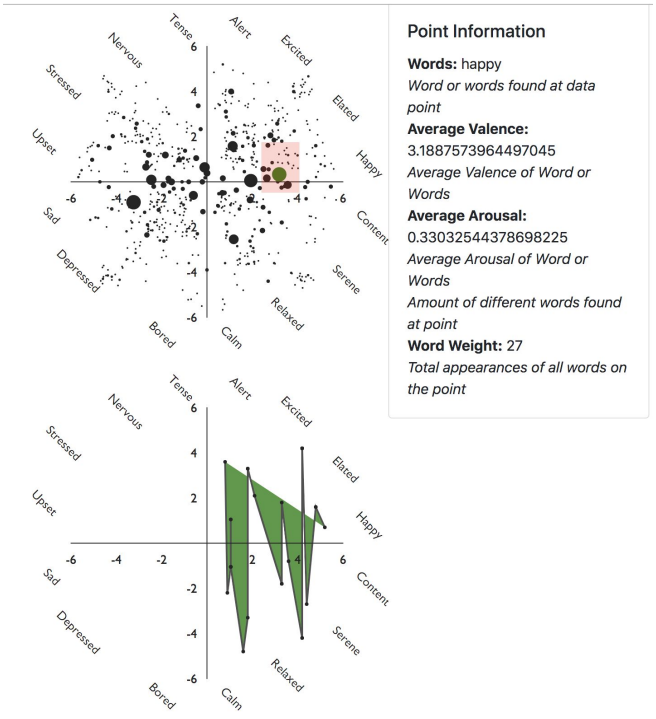
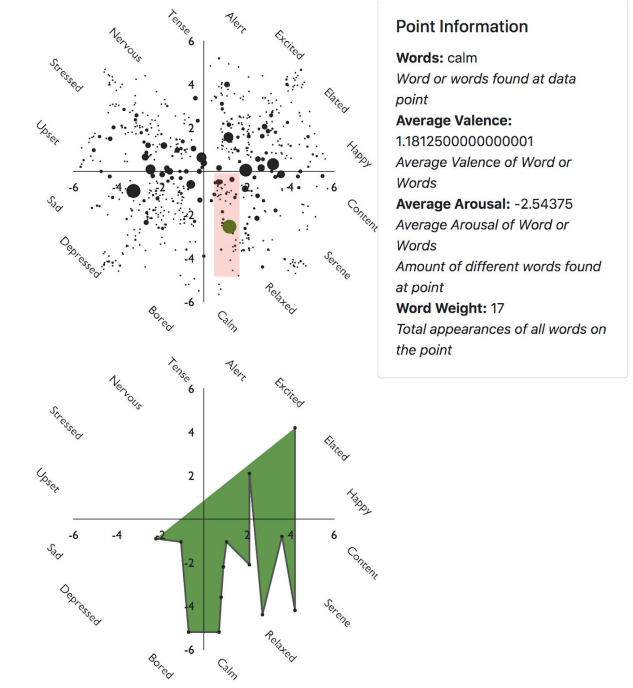


Figure 14: Word cloud from the Sensitive Pictures self-reported data.

The following examples show the emotional footprints of several other words from the Munch dataset, some of which (happy, calm and relaxed) correspond to recognised labels on the Circumplex model while others (fear) locate further concepts.





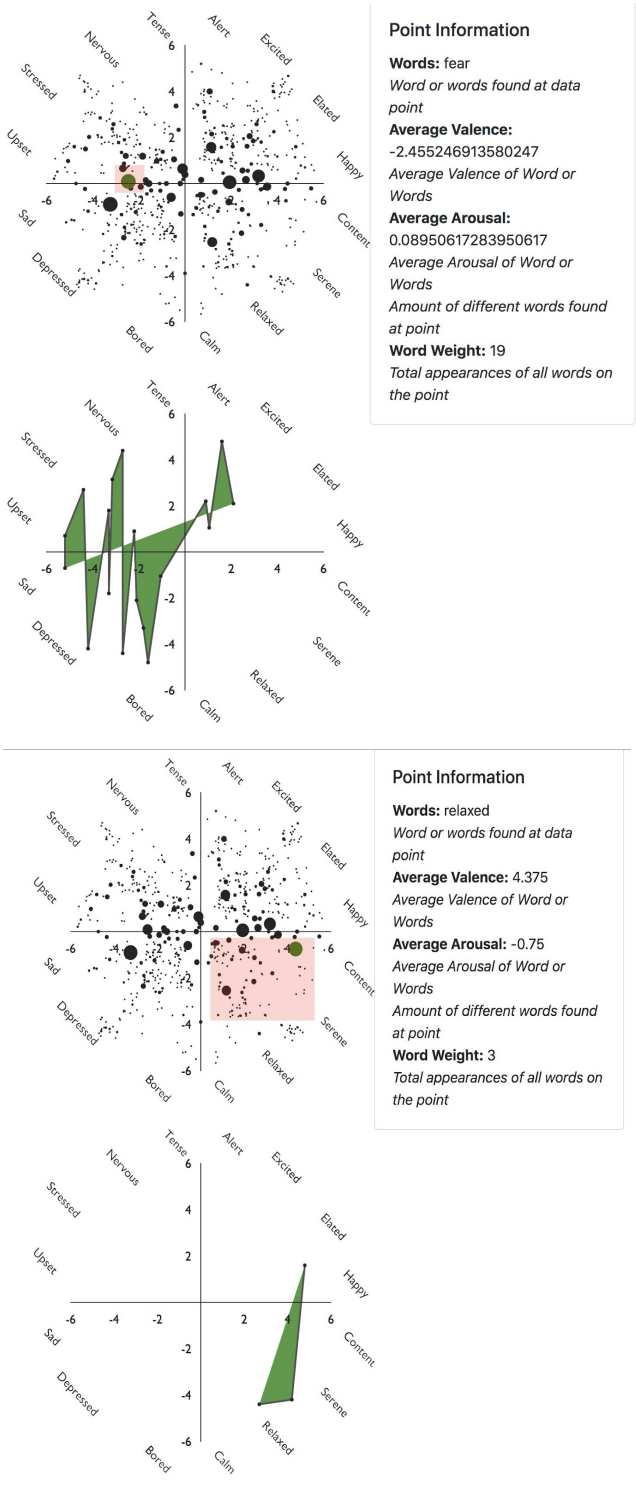


Figure 15: Emotional footprints of words from the self-reported data.

6 Sensitive Pictures with EEG headsets

6.1 EEG prototype

Based on insights from the August 2019 test, a new version of the prototype was developed, as described in deliverable D3.3. This prototype uses electroencephalogram (EEG) technology with sensors integrated into headphones used for listening to the museum audio guide. Thus the application can gather emotion data in real time during the audio experience. In this prototype the main web app works the same way as in the previous version, but there is no separated “kiosk” experience at the end. Instead, the final video is displayed in the web app, and is selected on the basis of the EEG data.

Regarding the technological setup, the system collects the EEG signal from 11 measuring locations on the scalp, plus the reference (REF) and ground (DRL). There are 4 electrodes around each ear, which makes 8 electrodes in total (relying on the principles of the ceegrid system - www.ceegrid.com), and 3 midline electrodes (C3, C4, and Cz, according to the 10-20 system). Headphones stream the raw signal (collected at 500Hz, 133Hz flat frequency response) to the user’s mobile device. To track the valence and arousal values, 1s windows, overlapped by 0.5s, were used. The valence was measured as an alpha asymmetry, whereas the arousal level was measured as a theta/beta ratio on the Cz electrode.

The measurement session starts when the subject presses the button to listen to the Sensitive Pictures narrative on their mobile phone. The “start" trigger is sent from the web portal to the EEG recording app, and the data collection starts. The valence and arousal levels are extracted every 0.5 seconds. Once the narrative ends, the portal sends the “end" trigger to the recording mobile phone. The arousal and valence data are projected onto the circumplex model valence-arousal plain. For each session and each participant, 1 value is provided for each of the 8 sections of the valence-arousal circle. This value represents the portion of time that the valence-arousal value was found in this particular section of the valence-arousal circle.

6.2 Data visualisations from the EEG prototype

The EEG prototype was offered to museum visitors during 21-23 November 2019. 60 participants tried the prototype in this period, of which 4 were museum employees and 56 were regular visitors. Furthermore, the prototype was demoed at the GIFT conference event in Lisbon on 27 November, at which time 29 participants tried the prototype. Note however that on this occasion participants were offered the choice to use regular headphones rather than the EEG sensing ones, and 14 of the test participants chose this option.

In the following, we present visualisations of data from the EEG prototype similar to the ones presented earlier. The visualisations display the self-reported values and EEG data for each of the six paintings that were part of the experience.

Each circle is positioned relative to the self-reported values from the participants, positioned on the circumplex model of emotions, and can contain multiple participant responses. Afterwards, each point has the EEG values displayed as up to 8 circles. The size and opacity of each circle is directed by the intensity value of each K value inside the EEG data, the larger and more opaque circles denouncing higher intensity. Hovering over each circle will display the amount of participant responses included inside the point, along with the free text that was input.

Each colour displayed on the data point represents a different K value, along with an associated emotion. These are Pleasant / Green, Excitement / Red, Neutral High / Pink, Stress / Yellow, Unpleasant / Black, Depression / Grey, Neutral Low / White, and Relaxation / Blue.

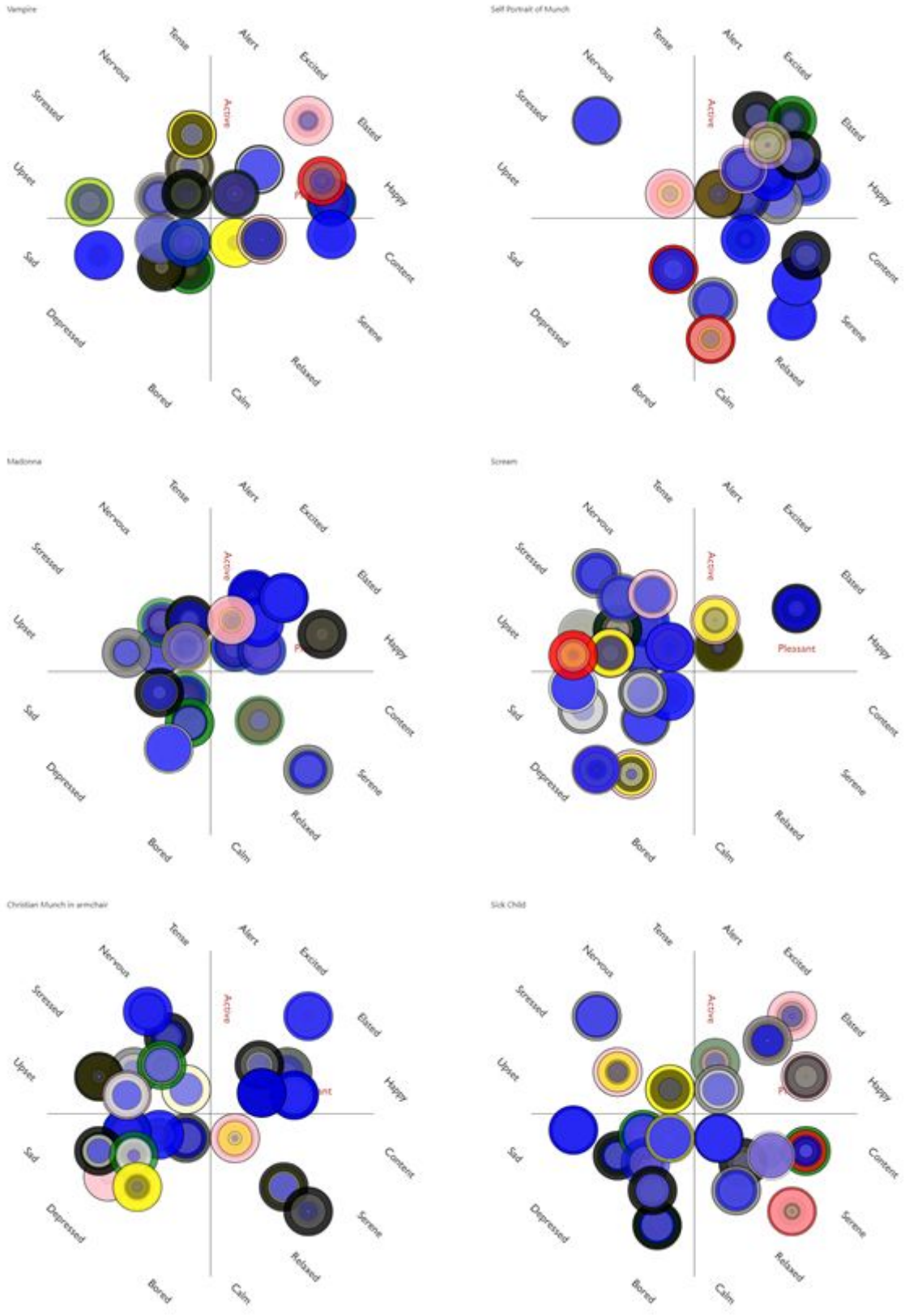


Figure 16: Six Exhibit profiles with EEG data.

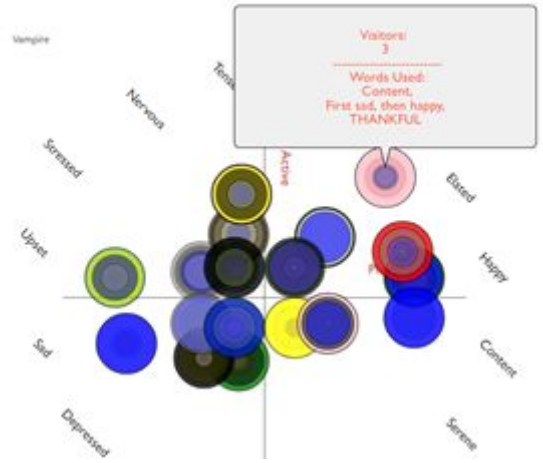


Figure 17: Mousing over a data point reveals the number of contributors and the words they used.

7 Concluding remarks

The reactions to the Sensitive Pictures experience from test participants indicate that most participants have found the experience engaging and meaningful, demonstrating that the prototype has potential to be developed into a successful product. The prototype that relied on visual emotion detection and that was tested in the Munch museum in August 2019 suffered somewhat from limitations connected with the computer vision technology; in particular that detecting emotions in this way require placing them in front of a camera, which was not possible to do inside the main exhibition space, and thus the emotion detection had to happen as a separate experience at the end of the visit. The use of EEG headsets in the final prototype removes this limitation, making it possible to gather computational emotion data during the main experience, and comparing them more directly to self-reported data, which increases the relevance and the value of the data.

Interviews demonstrate that our test users find the visualisations of their data interesting, and they appreciate the opportunity to reflect on their own emotions in relation to the artworks and the overall experience. We suggest that our approach offers the potential for future development in several key areas:

- **Enhancing museum experiences with emotional interpretation** - as a mechanism for encouraging emotional reflection during and after museum visiting experiences as noted above. Visitors might contribute their data at various points during an experience, and subsequently encounter the resulting visualisations and on signage, posters, interactive displays, postcards, badges, tickets and other points. We foresee potential for museums to create a variety of artefacts around the museum (emotional labels on exhibits or badged on visitors) that could encourage playful and reflective engagements with exhibits and emotions.

- **The museum as a site for emotional inquiry** - more generally, our approach might be used to communicate models of emotion from psychology and enable people to engage with them, explore their boundaries and challenge them. This is important as such as models are adopted by Artificial Intelligence, for example in the field of Affective Computing. There is an ongoing debate as to the validity of such models and indeed the whole idea of whether emotions can be reduced to a few simple parameters in such ways or rather, are complex multi-faceted concepts that are socially constructed and subject to ongoing interpretation. Museums and similar cultural institutions might play a key role in fostering this debate, using mechanisms such as this.

- **Enriching human-computer dialogue in conversational interfaces** - in the longer term such tools and their underlying datasets may suggest new ways of enriching the emotional

vocabulary of conversational agents, with wider implications to all manner of products and interactions and of shaping work in Affective Computing.

- **Applications beyond museums** - our approach might also be applied in other sectors, for example as part of new ways of understanding people's emotional responses in consumer science and marketing or perhaps even mental health and wellbeing.

References

Benford, S., & Giannachi, G. (2011). Performing Mixed Reality. The MIT Press.

Bennett, T. (1995). The Birth of the Museum: History, Theory, Politics. Routledge.

Bentley, T., Johnston, L., & von Baggo, K. (2005). Evaluation using cued-recall debrief to elicit information about a user's affective experiences. Citizens Online: Considerations for Today and the Future, 1–10. Canberra, Australia: ACM.

Betella, A., & Verschure, P. F. M. J. (2016). The Affective Slider: A Digital Self-Assessment Scale for the Measurement of Human Emotions. PLOS ONE, 11(2), e0148037. https://doi.org/10.1371/journal.pone.0148037

Boehner, K., DePaula, R., Dourish, P., & Sengers, P. (2005). Affect: From Information to Interaction. Between Sense and Sensibility. Presented at the 4th decennial conference on Critical computing, Aarhus, Denmark.

Boehner, K., DePaula, R., Dourish, P., & Sengers, P. (2007). How emotion is made and measured. International Journal of Human-Computer Studies, 65(4), 275–291.

Boehner, K., Sengers, P., & Gay, G. (2006). Affective presence in museums: Ambient systems for creative expression. Digital Creativity, 16(2), 79–89.

Damasio, A. R. (1995). Descartes' Error: Emotion, Reason, and the Human Brain. Harper Perennial.

Dodd J. and Sandell R. (2001) Including Museums: Perspectives on Museums, Galleries and Social Inclusion: Research Centre for Museums and Galleries, University of Leicester.

Duncan S.A. (2002) From Period Rooms to Public Trust: The Authority Debate and Art Museum Leadership in America. Curator: The Museum Journal 45: 93-108.

Eggum, A. (1984). Edvard Munch: Paintings, Sketches, and Studies. C.N. Potter.

Fosh, L., Benford, S., Reeves, S., Koleva, B., & Brundell, P. (2013). 'See Me, Feel Me, Touch Me, Hear Me': Trajectories and Interpretation in a Sculpture Garden. CHI 2013: Changing Perspectives. Presented at the SIGCHI Conference on Human Factors in Computing Systems, Paris, France.

Frank, M.G., Ekman, P. and Friesen, W.V., 1993. Behavioral markers and recognizability of the smile of enjoyment. Journal of personality and social psychology, 64(1), p.83.

Gaver, W. (2009, December). Designing for emotion (among other things). Philosophical Transactions of The Royal Society B Biological Sciences, 364(1535), 3597–3604.

Grafsgaard, J. F., Wiggins, J. B., Vail, A. K., Boyer, K. E., Wiebe, E. N., Lester, J. C. (2014). The additive value of multimodal features for predicting engagement, frustration, and learning during tutoring. In: Proceedings of the 16th ACM International Conference on Multimodal Interaction (pp. 42–49).

Hammal, Z., Cohn, J.F., Heike, C. and Speltz, M.L., 2015. Automatic measurement of head and facial movement for analysis and detection of infants' positive and negative affect. Frontiers in ICT, 2, p.21.

Höök, K. (2006). Designing familiar open surfaces. Fourth Nordic Conference on Human-Computer Interaction, 242–251.

Höök, K. (2012). Affective Computing. In The Encyclopedia of Human-Computer Interaction (2nd ed.). Interaction Design Foundation.

Höök, K., Ståhl, A., Sundström, P., & Laaksolahti, J. (2008). Interactional empowerment. ACM CHI 2008 Conference on Human Factors in Computing Systems, 647–656.

Isbister, K., Höök, K., Sharp, M., & Laaksolahti, J. (2006). The Sensual Evaluation Instrument: Developing an Affective Evaluation Tool. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1163–1172. https://doi.org/10.1145/1124772.1124946

Katz, J. (1999). How Emotions Work. University of Chicago Press.

Krippendorff, Klaus. Content Analysis : an Introduction to Its Methodology Fourth edition. Los Angeles: SAGE, 2019.

Krosnick, Jon A. (1991). "Response strategies for coping with the cognitive demands of attitude measures in surveys". Applied Cognitive Psychology. 5 (3): 213–236. doi:10.1002/acp.2350050305. ISSN 0888-4080

Ledoux, J. (1996). The Emotional Brain: The mysterious underpinnings of emotional life. Simon and Schuster.

Lynch B. (2013) Reflective debate, radical transparency and trust in the museum. Museum Management and Curatorship 28: 1-13.

Macnaghten P., Davies S.R. and Kearnes M. (2019) Understanding Public Responses to Emerging Technologies: A Narrative Approach. Journal of Environmental Policy & Planning 21: 504-518.

McCarthy, J., & Wright, P. (2007). Technology as Experience. The MIT Press.

McStay, A. (2018). Emotional AI: The Rise of Empathic Media. SAGE.

Norman, D. A. (2004). Emotional Design: Why We Love (Or Hate) Everyday Things. Basic Books.

Nzinga F. (2016) Public Trust and Art Museums. In: The Incluseum (ed). incluseum.com: The Incluseum.

Passebois J. and Aurier P. (2004) Building consumer/arts institution relationships: An exploratory study in contemporary art museums. International Review on Public and Nonprofit Marketing 1: 75-88.

Picard, R. (1997). Affective Computing. Cambridge: MIT Press.

Sanchez-Lozano, E., Tzimiropoulos, G., & Valstar, M. (2018). Joint Action Unit localisation and intensity estimation through heatmap regression. *British Machine Vision Conference*. Retrieved from http://arxiv.org/abs/1805.03487

Sharma, M., Kacker, S., & Sharma, M. (2016, December). A Brief Introduction and Review on Galvanic Skin Response. 2(6), 13–17.

Smith, L., Wetherell, M., & Campbell, G. (Eds.). (2018). Emotion, Affective Practices, and the Past in the Present. Routledge.

Sundström, P., Ståhl, A., & Höök, K. (2007). In situ informants exploring an emotional mobile messaging system in their everyday practice. International Journal of Human-Computer Studies, 65(4), 388–403.

Wetherell, M. (2015). Trends in the Turn to Affect: A Social Psychological Critique. Body & Society, 21(2), 139–166. https://doi.org/10.1177/1357034X14539020

Witcomb, A. (2014). 'Look, Listen and Feel': The first peoples exhibition at the Bunjilaka Gallery, Melbourne Museum. Thema La Revue Des Musées de La Civilistion, 1, 49–62.

Witcomb, A. (2015). Cultural pedagogies in the museum: Walking, listening and feeling. In M. Watkins, G. Noble, & C. Driscoll (Eds.), Cultural Pedagogies and Human Conduct (pp. 158–170). London, Eng.: Routledge.